

AD-A161 361

MODELING FAULT DIAGNOSIS PERFORMANCE ON A MARINE
POWERPLANT SIMULATOR(U) GEORGIA INST OF TECH ATLANTA
SCHOOL OF INDUSTRIAL AND SYSTEMS. Y D SU AUG 85

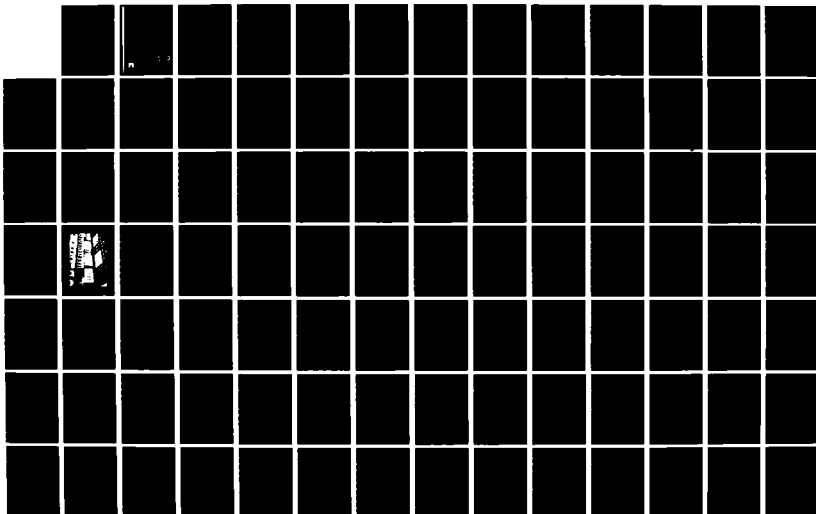
1/2

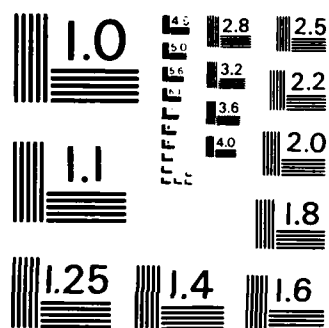
UNCLASSIFIED

GIT-TR-85-3 N00014-82-K-0487

F/G 13/10

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

6

center for man-machine systems research

**modeling fault diagnosis performance
on a marine powerplant simulator**

AD-A161 361

yuan-liang david su

report no. 85-3

august 1985

DTIC FILE COPY

**DTIC
ELECTE
NOV 14 1985
S E D**



**school of industrial and systems engineering
georgia institute of technology
a unit of the university system georgia
atlanta, georgia 30332**

**This document has been approved
for public release and sale; its
distribution is unlimited.**

85 11 12 08 9

AD-A161361

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION Unclassified		1b RESTRICTIVE MARKINGS	
2a SECURITY CLASSIFICATION AUTHORITY		3 DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited	
2b DECLASSIFICATION/DOWNGRADING SCHEDULE			
4 PERFORMING ORGANIZATION REPORT NUMBER(S) Technical Report 85-3		5 MONITORING ORGANIZATION REPORT NUMBER(S)	
6a. NAME OF PERFORMING ORGANIZATION Center for Man-Machine Systems Research	6b OFFICE SYMBOL (If applicable)	7a NAME OF MONITORING ORGANIZATION Personnel and Training Research Programs Office of Naval Research (Code 442PT)	
6c ADDRESS (City, State, and ZIP Code) School of Industrial & Systems Engineering Georgia Institute of Technology Atlanta, GA 30332		7b ADDRESS (City, State, and ZIP Code) 800 North Quincy St. Arlington, VA 22217-5000	
8a NAME OF FUNDING/SPONSORING ORGANIZATION Office of Naval Research	8b OFFICE SYMBOL (If applicable)	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-82-K-0487	
8c ADDRESS (City, State, and ZIP Code) 800 North Quincy St. Arlington, VA 22217-5000		10 SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO NR 154-491	
		PROJECT NO.	TASK NO.
		WORK UNIT ACCESSION NO.	
11 TITLE (Include Security Classification) Modeling Fault Diagnosis Performance on a Marine Powerplant Simulator			
12 PERSONAL AUTHOR(S) Yuan-Liang David Su			
13a TYPE OF REPORT Technical	13b TIME COVERED FROM 6/1/82 TO 5/31/86	14 DATE OF REPORT (Year, Month, Day) August 1985	15 PAGE COUNT 161
16 SUPPLEMENTARY NOTATION			
17 COSATI CODES FIELD GROUP SUB-GROUP		18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number) fault diagnosis; problem solving; training; simulators; performance modeling	
19 ABSTRACT (Continue on reverse if necessary and identify by block number) Human operators of large dynamic systems must possess good fault diagnosis skills. Factors affecting fault diagnosis skills are investigated in this thesis, and a model for fault diagnosis is developed. A low fidelity simulator of a marine powerplant was developed and used to study expert marine engineers' fault diagnosis performance. Based on the data collected, factors affecting the fault diagnosis performance were identified. They are: the initial feasible set (IFS) and the transition strategy. A subject's initial set of actions upon observing the symptoms forms the IFS which reflects the subjects' knowledge about a given failure. Transition strategy is concerned with the schemes subjects used to shift from the hypothesis formation stage to hypothesis evaluation stage. Two types of strategy were identified: breadth-depth strategy and balanced strategy. The data seem to indicate that subjects who had good IFS and used breadth-depth strategy performed better than subjects who had bad IFS and used balanced strategy. This finding seems to imply			
20 DISTRIBUTION/AVAILABILITY OF ABSTRACT <input type="checkbox"/> UNCLASSIFIED/UNLIMITED <input checked="" type="checkbox"/> SAME AS RPT <input type="checkbox"/> DTIC USERS		21 ABSTRACT SECURITY CLASSIFICATION Unclassified	
22a NAME OF RESPONSIBLE INDIVIDUAL Michael G. Shafto		22b TELEPHONE (Include Area Code) 202-696-4596	22c OFFICE SYMBOL

19. that both context-specific (IFS) and context-free (transition strategy) factors are important for fault diagnosis training.

A model for the fault diagnosis process which employs a conceptual entity called hypothesis frame to account for the observations from data and protocols has been proposed. Relevant information about a failure is "compiled" in the frames so that subjects could make quick and reliable inferences. Given symptoms of a failure, subjects would first try to match a frame. Once a frame is identified, subjects could use the information contained in the slots to make inferences. When no obvious frames could be processed, subjects could search in the system or symptom knowledge base for relevant information under the guidance of heuristics. This model could form the basis for implementing an on-line coach system for fault diagnosis.

MODELING FAULT DIAGNOSIS PERFORMANCE ON A MARINE POWERPLANT SIMULATOR

A THESIS

Presented to

The Faculty of the Division of Graduate Studies

By

Yuan-Liang David Su

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

in the School of Industrial and Systems Engineering

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

Georgia Institute of Technology

August, 1985

QUALITY
SELECTED

Approved for public release;
distribution unlimited

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	iv
LIST OF TABLES	v
LIST OF ILLUSTRATIONS	vi
Chapter	
I. INTRODUCTION	1
II. SIMULATOR FIDELITY	6
Overview of Simulator Training	
Definition of Fidelity	
Measurement of Fidelity	
Components of Fidelity	
Summary	
III. THE LOW FIDELITY SIMULATOR	28
Comparisons of Three Simulators	
The Low Fidelity Simulator -- FAIL	
Summary	
IV. DESIGN OF THE EXPERIMENTS	42
Subjects	
Experimental Procedures	
Experiments and Data Collection	
Summary	
V. ANALYSIS OF EXPERIMENTAL RESULTS	50
A Framework for Fault Diagnosis	
Factors Affecting Fault Diagnosis Performance	
Analysis of Experimental Results	
Summary and Conclusions	
VI. A MODEL OF FAULT DIAGNOSIS PERFORMANCE	76
Knowledge Representation	
Apparent Characteristics of	
the Diagnostic Process	
Hypothesis Frame	
A Model for the Diagnostic Process	
Interpretation of Protocols	
Conclusions	

VII. SUMMARY AND CONCLUSIONS	113
APPENDIX A List of Failures used in the pilot experiment	115
APPENDIX B List of Failures used in Sessions 1 to 5	117
APPENDIX C Protocols	130
APPENDIX D System Schematics	140
APPENDIX E Instruction Sheet	150
APPENDIX F Command List	152
REFERENCES	154
VITA	161

ACKNOWLEDGEMENTS

I would like to thank my advisor and friend, Dr. T. Govindaraj, for the guidance, help and patience he has provided throughout this research. A special thanks goes to Dr. William B. Rouse for his support and encouragement which began before and lasted throughout this research. I also wish to thank the committee members of my thesis, Drs. Christine Mitchell, William Johnson and Janet Kolodner for their time and substantial contributions to this work.

I would also like to thank my colleagues at the Center for Man-Machine Systems Research for their friendship and stimulating discussions: John Hammer, Richard Henneman, Mike Lewis, Janet Fath and Wan Yoon.

I would like to thank my wife, Chu-Yau, for her support of my pursuing Ph.D. degree. And finally, I would like to dedicate this thesis to my wonderful daughters, Joy and Jane, who have succeeded in delaying this thesis.

This research was supported by ONR under Contract N00014-82-K-0487, Work Unit NR 154-491.

LIST OF TABLES

Table	Page
3-1 Comparison of three simulators	33
5-1 Initial feasible sets for Failure no. 1-6	57
5-2 Performance statistics	64
5-3 Performance comparison between good IFS and bad IFS	67
5-4 Performance comparison between BD strategy and BL strategy	68
5-5 Performance comparison between good IFS and BD and all others	69
6-1 Examples of symptom-cause rules	81
6-2 The Boiler System Module	85
6-3 Protocols from subject S31 while solving Failure 19	89
6-4 Hypothesis frame for a condensate pump failure . . .	91
6-5 Types of heuristics used by subjects	96
6-6 Protocols from S33 while solving Failure 9	101
6-7 Protocols from S32 while solving Failure 23	104
6-8 Protocols from S34 while solving Failure 8	105
6-9 Occurrence of the attributes in the protocols . . .	111

LIST OF ILLUSTRATIONS

Figure	Page
2-1 The hypothetical relationship among fidelity, transfer and cost	12
2-2 The hypothetical relationship among degree of simulation and stage of learning	14
3-1 A view of the High Fidelity Simulator Console . . .	31
3-2 Main screen layout of FAIL	34
3-3 Fuel oil subsystem	36
5-1 Forward reasoning	51
5-2 Backward reasoning	51
5-3 A symptom-cause tree	53
5-4 Stages of diagnostic process	53
5-5 Performance based on total time spent and total number of actions	63
6-1 A hypothetical hierarchical tree	84
6-2 Relationships among known symptom set, symptom knowledge base, hypothesis frame, and system knowledge base	92
6-3 Flowchart of the diagnostic process	97
6-4 A simplified model of the air supply system which was implied by S33 in his protocol	102

Modeling Fault Diagnosis Performance
on a Marine Powerplant Simulator

Yuan-Liang D. Su

161 Pages

Directed by Dr. T. Govindaraj

Human operators of large dynamic systems must possess good fault diagnosis skills. Factors affecting fault diagnosis skills are investigated in this thesis, and a model for fault diagnosis is developed.

A low fidelity simulator of a marine powerplant, was developed and used to study expert marine engineers' fault diagnosis performance. Based on the data collected, factors affecting the fault diagnosis performance were identified. They are: the initial feasible set (IFS) and the transition strategy. A subject's initial set of actions upon observing the symptoms forms the IFS which reflects the subjects' knowledge about a given failure. Transition strategy is concerned with the schemes subjects used to shift from the hypothesis formation stage to hypothesis evaluation stage. Two types of strategy were identified: breadth-depth strategy and balanced strategy. The data seem to indicate that subjects who had good IFS and used breadth-depth strategy performed better than subjects who had bad IFS and used balanced strategy. This finding seems to imply that both context-specific (IFS) and context-free (transition strategy) factors are important for fault diagnosis training.

A model for the fault diagnosis process which employs a conceptual entity called hypothesis frame to account for the observations from

data and protocols has been proposed. Relevant information about a failure is "compiled" in the frames so that subjects could make quick and reliable inferences. Given symptoms of a failure, subjects would first try to match a frame. Once a frame is identified, subjects could use the information contained in the slots to make inferences. When no obvious frames could be processed, subjects could search in the system or symptom knowledge base for relevant information under the guidance of heuristics. This model could form the basis for implementing an on-line coach system for fault diagnosis.

CHAPTER I

INTRODUCTION

Humans encounter various types of problems everyday. Some of these difficult problems require the human to have advanced skills and training in order to be able to solve them. The quality of the solutions and the rapidity with which the problems are solved depend very much on the competence of the problem solvers. Therefore, it is interesting and useful to find out what type of skills a problem solver should possess to solve problems successfully.

Human problem solving has been studied from several different perspectives. Within the past few decades, mathematicians (Polya 1945), computer scientists (Newell and Simon 1972, Shortliffe 1976, Amarel 1983, Davis 1983), cognitive psychologists (Tversky and Kahneman 1975, Anderson 1980) and system engineers (Duncan and Shepherd 1975, Rouse 1981, 1983, Rouse and Hunt 1984) have contributed substantially to understanding of human problem solving. Due to the complexity of realistic problems, most of the problems studied in a controlled laboratory environment were simple game-type problems. There are many remaining questions regarding how results from this previous research can be applied to real problem solving environments such as troubleshooting in a large complex engineering system.

The human operator's role in the operation of large complex systems has become increasingly supervisory and managerial, requiring higher level perceptual and cognitive skills. Operator intervention

occurs when an undesirable or unforeseen event requires actions to detect, diagnose and compensate for failures, or dynamic conditions. Therefore, the operation of a large system sometimes depends on the problem solving skills of the operators. This type of problem solving is usually referred to as fault diagnosis in the literature (Duncan and Shepherd 1975, Rouse and Hunt 1984).

Fault diagnosis is a complex activity, requiring a considerable amount of practice and experience before one becomes proficient. Simulators are normally used to train people to improve their fault diagnosis performance. One objective of this research is to identify the skills required and factors affecting fault diagnosis performance, so that training simulators can be designed for specific operational environments.

In addition to training on actual systems or simulators, fault diagnosis performance of operators can be improved considerably by providing advice or aid at appropriate times. An effective computer aid can function as an on-line coach for troubleshooting by monitoring system states and processing information just like a human operator would in a fault diagnosis situation. A good model of human operator fault diagnosis performance is necessary to design an aid. The model must incorporate all relevant symptom, system and failure knowledge into a coherent structure. This model must also be able to reason and diagnose like a domain expert. A major problem in developing such a model is to determine what types of knowledge should be included and how they should be represented. In this thesis a model for fault diagnosis is proposed

as a preliminary step for building of an on-line coach.

In order to achieve the objectives discussed above, problem solving performance of marine engineer trainees was studied via a series of experiments conducted at Marine Safety International (MSI), in New York. The trainees, who were subjects in our experiments, were participants in a training course designed to teach marine engineers with 10 to 20 years of experience how to diagnose and compensate for powerplant failures in a steam-powered ship. During the training, realistic powerplant failures compiled by expert marine engineers were used on simulators of different fidelity levels. Their problem solving performance was carefully monitored. Factors affecting fault diagnosis are identified and tested. Also a process model of fault diagnosis is proposed and used to interpret protocols collected.

Simulators of a marine powerplant were used to create a realistic atmosphere so that "real" fault diagnosis behavior were observed. If a simulator resembles a real ship in providing an environment for failures, the behavior of trainees should be reasonably realistic. However, resemblance or realism is not easy to measure. Conventionally, transfer of training experiments are conducted to achieve this purpose. But transfer experiments are difficult, expensive and sometimes impossible to conduct (Adams 1979, Duncan and Shepherd 1975). Therefore, instead of conducting a transfer of training study, we used a descriptive approach. We analyzed the factors affecting the "realism" or "simulator fidelity" and investigated how simulator fidelity interacts with training effectiveness. The key factors that affect training

effectiveness were incorporated into the simulators. Chapter II describes this effort. An in-depth survey of simulator fidelity is conducted, the definitions, relationship to training and the components of fidelity are discussed.

Based on these discussions, a definition of fidelity was adopted. Simulators of high, moderate and low fidelity levels are described in Chapter III. It is then argued that, for training of cognitive behavior such as fault diagnosis, a low fidelity simulator is reasonably adequate. The description of the low fidelity simulator then follows.

This low fidelity simulator was used as a vehicle to study the fault diagnosis behavior of experts on marine powerplant failures. Experiments were conducted using this simulator. Chapter IV describes the experiments, including a description of the subjects, the experimental procedures, and the data collection technique.

Chapter V proposes a framework of fault diagnosis based on observations from experiments. Factors affecting fault diagnosis in the context of the framework proposed are derived and verified. The relationship between these factors and the performance was found to be statistically significant.

Although experimental results were statistically significant, the individual differences in diagnosis performance were striking. In order to probe this problem further, the relationship between the structure of the system and symptom knowledge and the mechanism of the fault diagnostic process were investigated. In Chapter VI, a model which is

abstracted from data and protocols collected is then proposed to account for the relationship between the knowledge structures and the diagnostic process.

Finally Chapter VII summarizes the findings of this research and discusses the implications for training and simulator development.

CHAPTER II

SIMULATOR FIDELITY

Traditional approaches to training on actual equipment are becoming more and more prohibitive because of relatively high cost and their limited ability to be used for training on unusual or potentially catastrophic situations. Simulators are used to cope both with increasing costs and limitations on training effectiveness. Spangenberg (1976) discusses seven unique advantages of using simulators for training. Simulators can (1) provide immediate feedback, (2) increase the number of system malfunctions and emergencies to provide the trainee with experience which would be unavailable on actual equipment, (3) compress time so a complex sequence of tasks may be accomplished in the time it would take to run through only one or two tasks on the actual equipment, (4) vary the sequence of tasks to maximize training efficiency, (5) provide guidance and stimulus support to the trainee in the form of prompts and feedback, (6) vary the difficulty level to match the skill level of each individual trainee, and (7) provide the trainee with an overview from which the trainee may form an overall understanding of the whole situation. These advantages, in addition to the potential cost-effectiveness are the reasons why simulators have been widely used.

Simulators take various forms. These include, as examples, mock-ups, photographic mimics, video disc and computer simulations. Usually simulators are less expensive than real systems. However, large mock-ups like those used to train pilots are expensive. The cost of a simulator usually increases with the fidelity of the simulator even

though increased fidelity does not guarantee better training.

Overview of Simulator Training

A "simulator" is a device or a facility which represents a machine, system, or environment and its functions (Gerathewohl 1969). Simulators have been widely used to train operators for maintenance, normal operations, problem-solving, and decision making. Simulators have been constructed for a variety of applications. Clymer (1980) identified at least eight different types : aircraft, aerospace, marine, ground vehicle, traffic, process plant, power plant, and manufacturing plant.

There have been numerous research efforts to evaluate the effectiveness of simulation for training. The extent to which a given simulator facilitates the acquisition of appropriate skills by the trainees is characterized by "transfer of training" from the training devices to actual equipment. This is called the "training effect" or "training effectiveness".

Training simulators may consist of several subsystems which interact with each other. Since each subsystem may contain hundreds of indicators and gauges, it can be expensive to construct and run such a simulator. Therefore, the question of how to utilize simulators efficiently becomes important.

Conventionally, simulators have been employed in training with the assumption that higher fidelity produces a better transfer effect. However, research that contradicts this assumption has also been

reported during the past few years (Johnson 1981, Johnson & Rouse 1982, Trollip 1977). The rest of this chapter discusses the general issue of simulator fidelity. Its definition, relationship with training, measurement and components are described.

Definitions of Fidelity

"Fidelity" has been used widely and diversely in the simulator training community. Different people use the term with different meanings. Hays (1981) reviewed the literature and noted the diversity of meaning. He found that most researchers contrasted physical fidelity with non-physical fidelity. It is non-physical fidelity that attracted a variety of names and definitions. Functional fidelity, psychological fidelity, task fidelity and behavioral fidelity (Hays 1980) are among the names used. In general, most researchers agree that physical fidelity is not the only factor, nor the main factor, affecting training effectiveness. There is also general agreement that higher fidelity (assuming it can be measured) is not necessary for every aspect of every kind of training (Hays 1981).

There appears to be a lack of consensus in the research on simulator fidelity, and of an appropriate definition of what is meant by fidelity. After reviewing several attempts to define simulator fidelity, Hays (1981) proposed the following definition :

Training simulator fidelity is the degree of similarity between the training simulator and the equipment which is simulated. It is a two dimensional measurement of this simi-

larity in terms of: the physical characteristics of the training simulator, and the functional characteristics of the simulated equipment.

Rouse (1982) defined fidelity:

"the precision with which the simulator reproduces the appearance and behavior of the real equipment."

These two definitions are very similar. They emphasize that fidelity is a two dimensional concept. They also pointed out the measurement problems. Tasks and the responses of the trainees were not explicitly considered.

According to Kinkade and Wheaton (1972), the fidelity of a simulator consists of three different components: (1) equipment fidelity (2) environmental fidelity, and (3) psychological fidelity. Equipment fidelity is defined as the degree to which the simulator duplicates the appearance and "feel" of the real system. Environmental fidelity is concerned with the degree to which the simulator duplicates the sensory stimulation, e.g., dynamic motion cues, visual cues, etc. Psychological fidelity is simply the degree to which the trainee perceives the simulator as a duplicate of the real system. Equipment fidelity is actually what Hays defined as physical fidelity, while the environmental fidelity and the psychological fidelity together approximate his functional fidelity. However, psychological fidelity explicitly recognizes the role of the trainees' perception of fidelity.

Govindaraj (1983) proposed a three-dimensional approach in which

further descriptions and measurements along each dimension are discussed.

(1) Physical fidelity :

Physical fidelity is concerned with the variables presented and the forms they take as well as the environmental factors such as noise, vibration and thermal conditions. Techniques from syntactic pattern recognition are suggested to measure physical fidelity.

(2) Structural fidelity :

Structural fidelity refers to the relationships between subsystems. Level of abstraction, coupling of system states, and aggregation of subsystems are the primary concerns. Graph theoretic methods are suggested for measurement.

(3) Dynamic fidelity :

Dynamic fidelity refers to the evolution of system states over time and their presentation to trainees. Control theoretic methods are suggested for measurement.

This definition appears to be relatively comprehensive and especially useful for describing the fidelity of simulators of large complex systems such as power plant control rooms. Non-physical fidelity is decomposed into structural fidelity and dynamic fidelity. This provides a way to analyze and measure the functional aspects of a simulator.

Despite the rigorous attempts to define simulator fidelity, one must keep in mind that training effectiveness is the main concern. If

high fidelity does not imply high transfer of training, then fidelity is not a useful concept. As pointed out by Rouse (1982), the key issue in the use of simulators is the level of fidelity necessary to assure transfer of training from simulators to real equipment. The study of simulator fidelity can help clarify the following questions.

- 1). What are the variables affecting the feeling of realism?
- 2). What is learned and in what way?
- 3). Can criteria for simulator design be found? What should the fidelity level of a training simulator be to achieve a given goal?
- 4). What is the relationship between each dimension of fidelity and transfer of training? Or, does any meaningful relationship exist between these two?

A definition of fidelity is necessary if any further study of fidelity is anticipated. It may not be possible to have a general index of fidelity for design purposes. Nevertheless, an explicitly expressed and commonly accepted definition is required for comparison of fidelity between different simulators.

Relationship With Training

A hypothetical relationship among fidelity, transfer, and cost was proposed by R. B. Miller (1954) (see Figure 2-1). Very little empirical data have been collected to explore this relationship. According to Miller, an increase in the degree of simulator fidelity is accompanied by increases in both transfer of training and cost. The

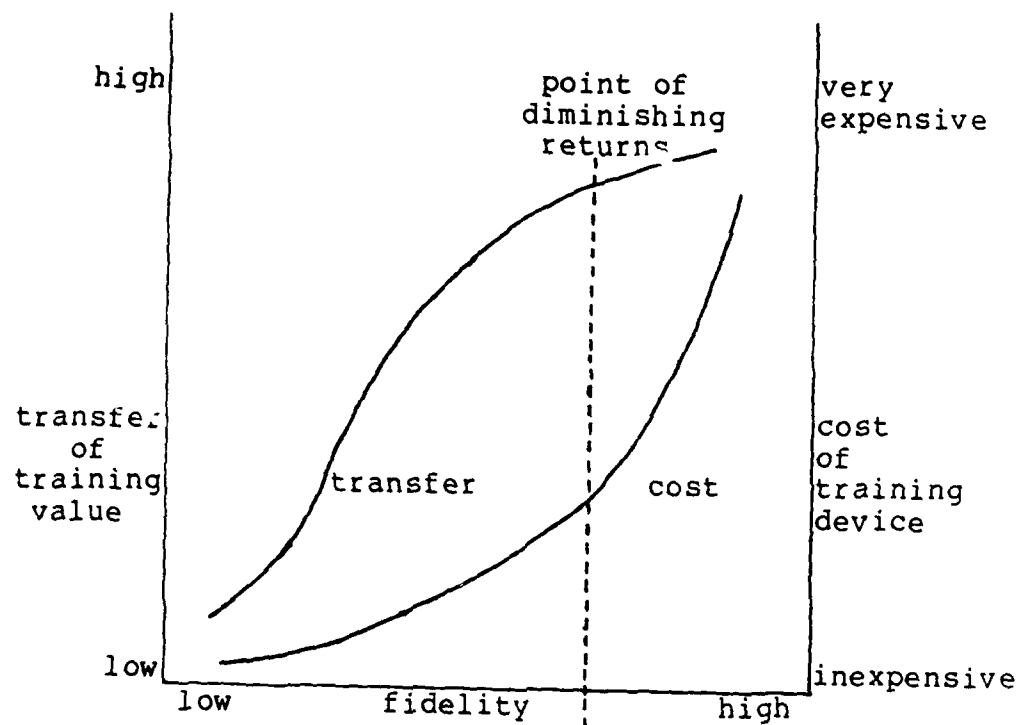


Figure 2-1 : The hypothetical relationship among fidelity, transfer and cost (modified from Miller 1954)

objective, both for simulator design and the use of a simulator for training, is to find the optimal point of intersection between fidelity, transfer and cost in each case. One problem with Miller's formulation is that the cost of a simulator could go to infinity as its fidelity

increases (Orlansky 1984). Another problem is the explicit assumption that the amount of transfer increases with increasing fidelity of the simulator (Micheli 1972). Many researchers have found that comparable training results may be obtained with both low- and high-fidelity simulators of the same equipment (Duncan and Shepherd 1975, Crawford and Crawford 1978, Johnson 1981). In a study by Martin and Waag (1978), it was shown that flight simulators with higher fidelity provided too much information for novice trainees and actually detracted from simulator effectiveness. Prophet (1966) reported a study that compared a low fidelity simulator (inexpensive photographic mock-up of a cockpit) with that of an elaborate trainer. No significant difference between groups was found. Despite these counterexamples, Miller's approach is cited widely (Fink and Shriver 1978, Kinkade and Wheaton 1972, Hays 1981).

A reformulation of Miller's view has been proposed by Orlansky (1984). Even though Orlansky's hypothetical model is not fully supported by empirical data, the known facts about the cost of simulators and about the relationship between transfer of training and fidelity have been accounted for in the model.

Kinkade and Wheaton (1972) have proposed a hypothetical relationship between the degree of simulator fidelity, types of simulator fidelity and the stages of learning (see Figure 2-2). Early in the training program (procedure training), the trainee does not benefit from high degrees of either physical or environmental fidelity. However, as skill is acquired (familiarization training), there are requirements for increases in both physical and environmental fidelity, with the require-

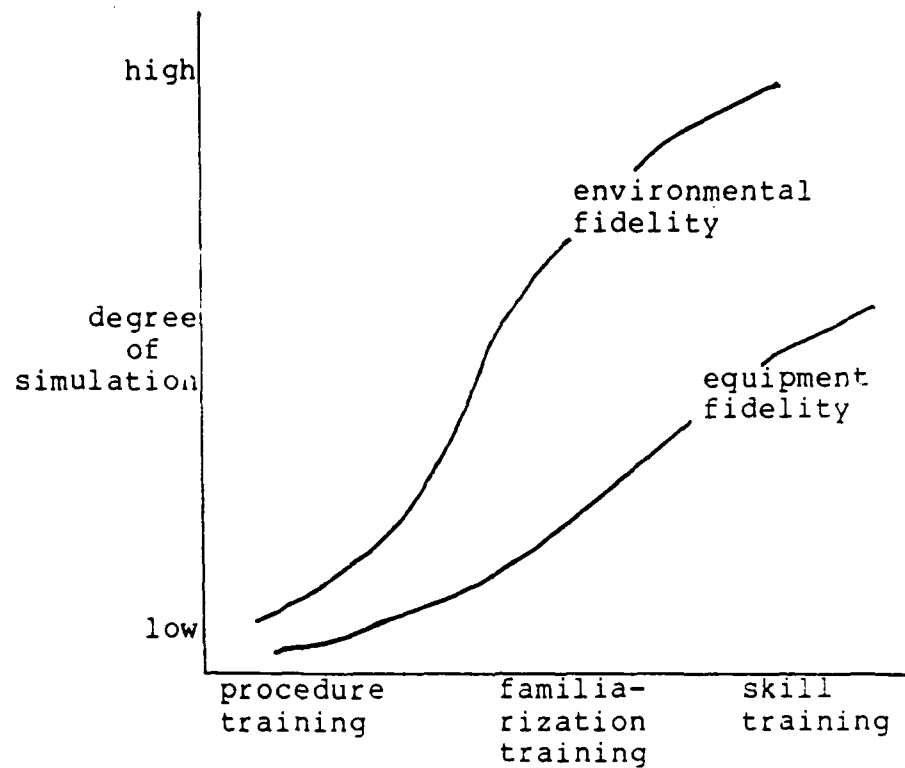


Figure 2-2 : The hypothetical relationship among degree of simulation (fidelity) and stage of learning (Kinkade and Wheaton, 1972)

ments for greater environmental fidelity increasing at a faster rate. During later stages of training (skill training), increases in both types of fidelity are desirable, with a requirement for higher levels of

functional fidelity.

Johnson (1981) was able to show that high fidelity is not required for training in procedural tasks. Several other researchers (Johnson and Rouse 1982, Johnson and Fath 1984, Maddox, Johnson and Frey 1985) reported similar results for fault diagnosis tasks. Govindaraj (1983) also casts doubt on the necessity of high physical fidelity for problem solving training. Baum (1981) pointed out that empirical data to support Kinkade and Wheaton's conjecture are lacking except those for procedure training. Baum, Riedel and Hays (1982) conducted a study to determine the relationship between training device fidelity and transfer of training for a perceptual-motor maintenance task. The results indicate that physical similarity is a significantly more important determinant of skill acquisition than functional similarity. These experiments provide some support for Kinkade and Wheaton's proposal.

Fink and Shriver (1978) made a point similar to that made by Kinkade and Wheaton. They identified four training stages: (1) acquisition of enabling skills and knowledge (2) acquisition of uncoordinated skills and unapplied knowledge (3) acquisition of coordinated skills and ability to apply knowledge and (4) acquisition of job proficiency. They claimed that different stages require different levels of fidelity with the first stage requiring the lowest level.

G.G. Miller (1974) drew the following conclusions about the relationship between fidelity and training.

- (1) High fidelity is never associated with poor training.
- (2) Transfer of training is more a function of how the

simulator is used rather than the degree of fidelity.

(3) Procedural task training does not require high fidelity.

Conclusions 2 and 3 are shared by many other researchers (Duncan and Shepherd 1975, Johnson 1981, Johnson and Rouse 1982), while Conclusion 1, as pointed out before, is doubtful.

No consensus has been reached on the relationships between fidelity and other factors such as cost, training, and stage of learning. The research in this area is not very conclusive. The difficulty of measuring fidelity is part of the reason for the slow progress. The next section discusses the problems and the alternatives for the measurement of fidelity.

Measurement of Fidelity

The measurement of fidelity is an important step if one wishes to determine empirically the relation between level of fidelity and training effectiveness as well as the necessary fidelity level of a simulator for training for a given task. Specific transfer of training studies are possible only after both the simulator and the actual equipment have been built. Nevertheless, there is a need to be able to predict the effectiveness of the training device prior to construction. Considering the tremendous cost and man-hours involved in developing simulators of any fidelity level, one cannot be satisfied with a post hoc measure. A measure of fidelity that correlates with the measure of transfer of training is a useful system design guide. Therefore, the purpose of measuring fidelity is the hope that a predictive index can be devised for anticipating the effectiveness of a training simulator. A reliable,

predictive index of the effectiveness of a simulator will be very useful both for trainers and design engineers. Other things being equal, such as user acceptance and required levels of funding, they can then choose only those features that possess high transfer value and still meet the training objective. However, in practice this is very hard to achieve due to the difficulty of measuring simulator fidelity. One of the difficulties is the lack of generality of such a measure. Govindaraj (1983) pointed out that the environment and the purpose for which the simulator is to be used have a strong influence on fidelity. Also, fidelity appears to be very context-specific. Therefore, it may be difficult to derive context-free measures of fidelity.

Wheaton et al. (1976) assessed simulator fidelity on two dimensions: physical fidelity and functional fidelity. They discussed the metrics of fidelity in the context of constructing a model to predict training device effectiveness. In their approach, a thorough task analysis of the target system and the simulator was conducted. Subtasks of the target system and the simulator were then clearly identified. The physical fidelity, for each subtask, between the real system and the simulator was evaluated by rating with a scale that ranged from "no resemblance", "dissimilar", "similar" to "identical". The functional fidelity was evaluated by recording the operator's behavior in terms of the information flow from each display to the operator, and from the operator to each control. For each subtask, the type, amount, and direction of information was assessed using information-theoretic methods. Then a four-point scale was applied by comparing the information metrics between the real system and the simulator on each subtask.

The underlying assumption was that the higher the rating on the assessment factors, the higher the transfer that would take place and the more effective the simulator. However, as pointed out by Adams (1979), rating is very subjective and its reliability is questionable. Further refinement of this assessment process was reported by Narva (1977), in which the physical fidelity and the functional fidelity were measured by rating with emphasis on behavioral categories instead of the original subtasks. Some of the behavioral categories used include rule learning and use, detection, symbol identification, decision making, etc.

Caro (1970) advocated a procedure called Equipment Device Task Commonality Analysis in which the measurement of fidelity was conducted by assessing the similarity of S-R relationships in the real system and the simulator. Positive transfer was assumed to occur when both stimuli and responses were similar. Negative transfer was predicted when the stimuli were similar but the responses were different. This is similar to what Osgood (1949) proposed. The assessment of the similarity was also accomplished through rating. This procedure applies only to simulators where the stimuli and the responses can be clearly identified. In a complex system, it may be impossible to specify the stimuli clearly.

Components of Fidelity

As pointed out in the previous discussion, "fidelity" is a mutli-dimensional concept. An operational, comprehensive definition may be difficult to obtain. However, the building blocks of fidelity have been widely noted and studied for a long time. These are the design features of a simulator. Some of them are discussed below. This list is definitely not exhaustive.

Stress

There is no doubt that stress is experienced by most operators of any real system. However, as Duncan and Shepherd (1975) pointed out, it is not clear how or if stress can be simulated during training. There are at least three types of stress. First, there is the feeling of danger. Creating this type of stress on a simulator during training is very difficult. Second, there is the threat of hazard or sanction. This form of stress can only be simulated by manipulating reward as a consequence of performance. Third, there is time stress. This can easily be introduced into the training task, but may alter the trainee's perception of the task. Not much is known about how to incorporate stress into simulator training or if its presence contributes to adequate training. While such an assessment is quite subjective, anyone who has spent a good deal of time in a flight simulator, or nuclear powerplant simulator will attest to the high stress feeling.

Environment

Noise is distracting especially in complex tasks that require close attention and concentration (Finkelman 1975). Improper lighting (Tinker 1943), temperature (Pepler 1972), etc. degrade human performance. However, how much these affect the fidelity level or how much they contribute to the training effects is a matter difficult to estimate. While noise, inappropriate lighting, and temperature may degrade general performance, systematic noise or unusual heat or temperature are repeatedly reported to be of great help for failure detection and diagnosis. Many trainees, designers and experienced operators admit the possibility of using unusual environmental changes as a clue to detect or diagnose the failure. Vibration has been given the same appraisal (Longman, Phelan and Hansford 1981, McCallum and Rawson 1981, Jaspers and Hanley 1980, Semple et al. 1981, Martin and Wagg 1978).

Layout

Panel layout, display size and even the coloring of instruments are considered to be important factors that affect the feeling of realism. More importantly is the relative distance and the relative position between gauges, annunciators and status indicators (Fowler et al. 1968). Duncan and Shepherd (1975) argued that the trainees may develop strategies that heavily depend on patterns of the presented stimuli. The size of the display may influence the amount of information the trainee can process at any one time. The relative distance between gauges and the relative position of stimuli may affect the pattern recognition process. However, Duncan and Shepherd pointed out that the

influence of such factors is unknown.

Wholeness

A full-scale simulator provides all aspects of system training, while a part-task simulator presents only selected parts of the full system to the trainees. The benefit of a part-task trainer is that some particularly important subsystem such as the turbine or the boiler may be represented with greater physical fidelity and provided for training before coping with the entire system. However, the functional fidelity may be affected due to the isolation of a particular subsystem. Curry (1981) observed that detection, diagnosis and remedial action are generally assumed to be three separate tasks. Therefore, training on each one can be accomplished independently without too much trouble. Rouse (1981) found that a particular logical judgement process is especially important for effective fault diagnosis. Abstracting this logical process, he developed a context-free task, TASK, which is, in some sense, a decomposition of the fault diagnosis behavior. He demonstrated positive transfer of training from TASK to a real system. Rasmussen (1980) proposed a criterion for the decomposition of a complex function. He observed that:

"...break-down of complex functions is only acceptable if the performance is paced by the system, i.e., cues from the system serve to initiate elementary, skilled sub-routines individually and to control their sequence. This is the case in many manual tasks, e.g., mechanical assembly, but can probably also be arranged in more complex mental tasks by properly designed

interface systems." (p. 92)

The influence of the part-task trainer on complex mental tasks, such as fault diagnosis and problem solving, is not yet clearly understood. However, the unverified conjecture is that wholeness is not a crucial fidelity factor.

Dynamics

Most real systems are dynamic, as are most simulators. However, static simulators have been used increasingly in the past few years (Duncan and Shepherd 1975, Shepherd et al. 1977, Hunt and Rouse 1981, Johnson and Rouse 1982). Static simulators only allow the operators to check the system status, while dynamic simulators accept control commands and execute them. There is no doubt that dynamic simulators describe the object task better than static simulators do, but how much better is a question unanswered. Forbus and Stevens (1981) indicated that there is a growing amount of evidence that human understanding of physical systems is based on qualitative models of those systems. This evidence comes from psychological studies (Larkin et al. 1980) and is supported by success in artificial intelligence in actually constructing systems that reason about physical situations using qualitative models (deKleer 1979, Forbus 1980). Govindaraj (1983) proposed a qualitative approach to modeling a complex dynamic system. This approach may provide a way to associate the level of dynamic fidelity with an explicit training effect. However, there is no empirical data to support the transfer effect of the qualitative dynamic simulator.

Abstraction

A physical system can be represented mentally in different forms (Rasmussen 1979, Rasmussen 1985). Simulators may be constructed to represent the physical system at different levels of abstraction. On the bottom of the hierarchy is the realization of the physical components in detail, analogous to a system mock-up. The higher the model stands in the hierarchy by aggregating elements into larger units or by abstracting through functional properties, the less the physical fidelity becomes. A system block diagram is an example of a more abstract simulator. Each level of abstraction possesses its own set of symbols and syntactic rules. Abstract simulators may be more effective in training for fault diagnosis due to the absence of irrelevant cues. Rasmussen argued that shifting between levels of abstraction for suitable strategy may be helpful for problem solving. This implies that training under lower physical fidelity and higher abstraction level may transfer well to higher physical fidelity and lower abstraction situation. The fact that diagnosis can be viewed as a top-down process may explain why lower physical fidelity and higher abstraction level simulators could perform better in this type of training. Therefore, functionally speaking, it is hard to decide which one has higher fidelity.

The value of simulators of different abstraction levels may be different for different levels of trainees. Kriessman (1981) speculated that simulators of different fidelity level may achieve the training effect differently. A high fidelity simulator is good for more experienced trainees, while a low fidelity simulator is better for less

experienced ones. However, it is still an open question as to whether the use of simulators of different abstraction levels may provide the operator with different skills or the same type of skills but in a degraded mode.

State Variables

Most of the state variables in a real system are presented in a continuous manner via gauges and meters, while for simplification, some simulators may represent the state variables in discrete language such as high/medium/low or on/off. Internally, the human seems to process information in a discrete manner (Kuipers and Kassirer 1984), especially when logical reasoning is involved. He may classify information into several finite sets. Presenting information in a discrete manner may not result in a loss of information as long as the classifying scheme matches the human's internal model.

The increasing use of CRTs for display in simulators introduces difficulty in presentation of state variables because of size constraints. The most common strategy is to use serial presentation instead of parallel presentation which is the usual way information is transferred to the operator in a real system. However, considering the human as a limited information processor, this restriction may not be as serious a fidelity problem as it first appears. The attention span for human beings is well known to be narrow and varying in time. The state variables in a real system, though presented simultaneously, are possibly processed in a serial manner --- perhaps chunk by chunk. However, how serial presentation of state variables affect fidelity may depend on

the type of simulators used.

Summary

Several attempts have been made to define "fidelity". Hays (1981) proposed a functional fidelity vs. physical fidelity approach. Rouse (1982) suggested a similar idea. Govindaraj (1983), oriented toward an operational definition, decomposed functional fidelity into structural fidelity and dynamic fidelity. Lack of empirical studies of fidelity issues makes it difficult to develop a useful definition of fidelity. A generally accepted assertion is that higher fidelity does not guarantee better transfer. Kinkade and Wheaton (1971) conjectured that the fidelity requirement varies with the stages of learning. Generally, it is proposed that procedural tasks do not require as high a fidelity as visual-motor skills do.

Wheaton (1976) and Narva (1977) developed a predictive index for transfer effect based on the measurement of fidelity. Task analysis of both the real system and the simulator is the foundation for measurement. Rating, so far, is employed in almost every fidelity metric. A more objective metric based on system characteristics is an important future research topic. Several factors that affect fidelity were also discussed. A brief summary is given below.

- (1) It is very difficult to include stress in the simulator.
- (2) Environmental factors such as noise, lighting, temperature, motion and vibration are annoying but may be treated as diagnostic aids. Inclusion of these variables does increase

fidelity, but the cost-effectiveness of including them in a simulator has long been challenged.

- (3) Layout may affect the strategy used by trainees.
- (4) The important issue in the use of part-task simulators is the decomposability of the tasks.
- (5) Dynamic features may not be crucial in training for fault diagnosis. Several studies indicated that the human reasons in a qualitative rather than quantitative way.
- (6) It may be beneficial to vary the level of abstraction of the simulator depending upon the level of skill of the trainee.
- (7) In a real system, the state variables are presented simultaneously, although humans may not be able to process all of this information at once. As a limited information processor, human operators may do well with serial presentation of the state variables.

In general, physical fidelity may be important for training for normal operations while non-physical fidelity plays a more important role in training for troubleshooting, especially those involving cognitive skills (Johnson 1981, Johnson and Rouse 1982). Following this argument, we use a low fidelity simulator of a marine powerplant in our research to study fault diagnosis behavior. This low fidelity simulator is in a family of simulators with increasing fidelity levels (low, moderate and high). The low fidelity simulator has moderate structural fidelity and low fidelity on both physical and dynamic fidelity according to the definition of fidelity used by Govindaraj (1983). In the next chapter the low fidelity simulator is described in detail,

following a comparison of the low, moderate, and high fidelity simulators.

CHAPTER III

THE LOW FIDELITY SIMULATOR

Simulators can be designed at various levels of complexity and realism relative to the real systems. The degree of realism and closeness with which a simulator resembles the actual system can be formalized in terms of the concept of fidelity of simulation. The issue of defining simulator fidelity was discussed in Chapter II; the difficulty of defining fidelity and the components of fidelity were considered in detail. A working definition of fidelity was proposed and used to classify simulators (Govindaraj 1983).

Fidelity was treated as a three-dimensional concept: physical, structural and dynamic fidelity. Physical fidelity is concerned with the variables that are displayed and in how much detail they are given. This includes environmental factors such as noise, vibration and thermal conditions. Structural fidelity refers to the nature of relationships between various subsystems that make up the system, including feedback and feedforward connections, and hierarchical relationships. Similarly, the evolution of the system states over time, and their representation can be characterized by dynamic fidelity.

Comparisons of Three Simulators

To understand how simulator fidelity affects human fault diagnosis behavior, experiments are needed using simulators at different fidelity levels. This chapter considers oil-fired marine powerplant

simulators at high, moderate and low fidelity levels based on the definition presented above.

A high fidelity simulator (HFS) is available at Marine Safety International (MSI), New York. Partial view of the simulator is shown in Figure 3-1. This simulator has two boilers of equal size with two burners per boiler. Two feedpumps supply water to the boilers. There is a high pressure turbine and a low pressure turbine. The turbines are controlled by a throttle control system at the control console. Behind the control room where the control console is housed, the engine room is simulated using components such as boilers, feedpumps, and vacuum pumps. The environment is simulated with appropriate noises, humidity, and temperature. The entire simulator is run by a PDP 11/70 minicomputer.

Gauges, alarms and status indicators are displayed on the control console. There are gauges for temperatures, pressures and levels of the boilers and other subsystems such as the condenser system. Status indicators can be grouped according to their functions. There are 12 groups: throttle control, main turbine bearing temperatures, STBD boiler burner management, PORT boiler burner management, salinity, STG bearing temperatures, bleed status, lube oil, condenser, turbine vias panel, fuel oil and feedpump systems. Six panels of alarms are shown in the middle section of the console. Each panel contains a number of indicators which are lighted as necessary to display the states of components and/or subsystems. They are alarms for auxiliary, STBD boiler, PORT boiler, feedwater, throttle control and vacuum systems. There are roughly 100 gauges, 200 alarms and 200 status indicators in the entire

system. This simulator has high fidelity along all three dimensions. Besides the apparent physical, dynamic and structural fidelity, schematics that describe the structure of the system are also available.

A moderate fidelity simulator is being developed at Georgia Tech (Govindaraj 1984) in which the dynamics of major subsystems are represented qualitatively. Even though the physical fidelity of this simulator may be low, fairly high structural fidelity and moderate dynamic fidelity are achieved. A low fidelity simulator called FAIL (FAULT-based Aid for Instruction and Learning) (the term FAULT is described in the next section) has been developed that runs on an APPLE II computer. System information is displayed to the operator upon request. Schematic diagrams of important subsystems are provided on paper. Even though the physical and dynamic fidelity are very low, the structural fidelity is close to the high and moderate fidelity simulators. Comparison of these three simulators in terms of the three dimensions of fidelity is summarized in Table 3-1.

Although HFS has high fidelity along all three dimensions, the high cost of construction and operation and the difficulties in data collection have prevented it from being used as a vehicle to study fault diagnosis behavior. As discussed in Chapter II, high non-physical fidelity may be sufficient for training of cognitive or procedure type of tasks. In this thesis the fault diagnosis behavior on the low fidelity simulator, FAIL, which has relatively high structural fidelity is described. The details of FAIL experiments in which experienced marine engineers participated as subject, are described next.

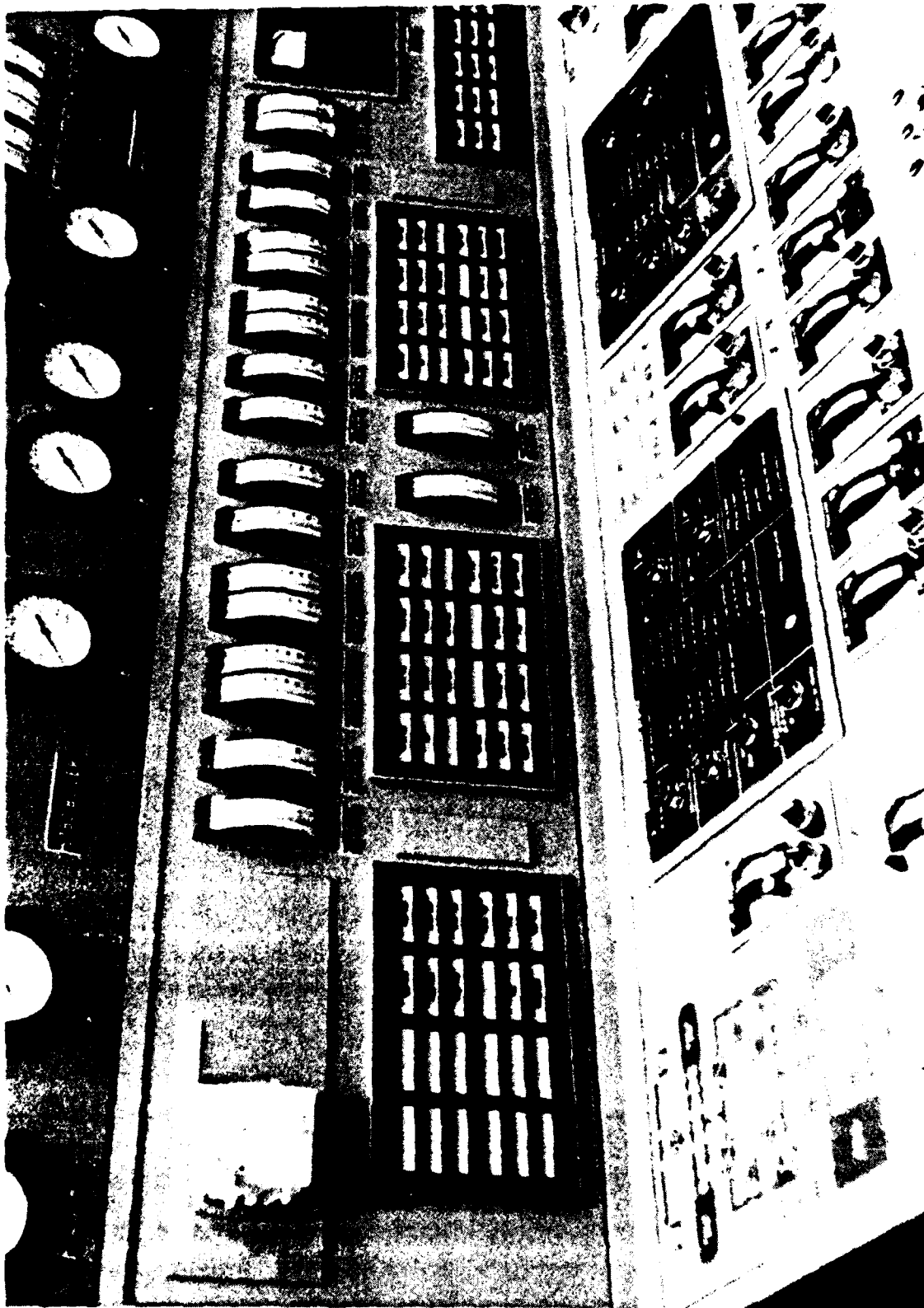


Figure 3-1: A view of the High Fidelity Simulator Console

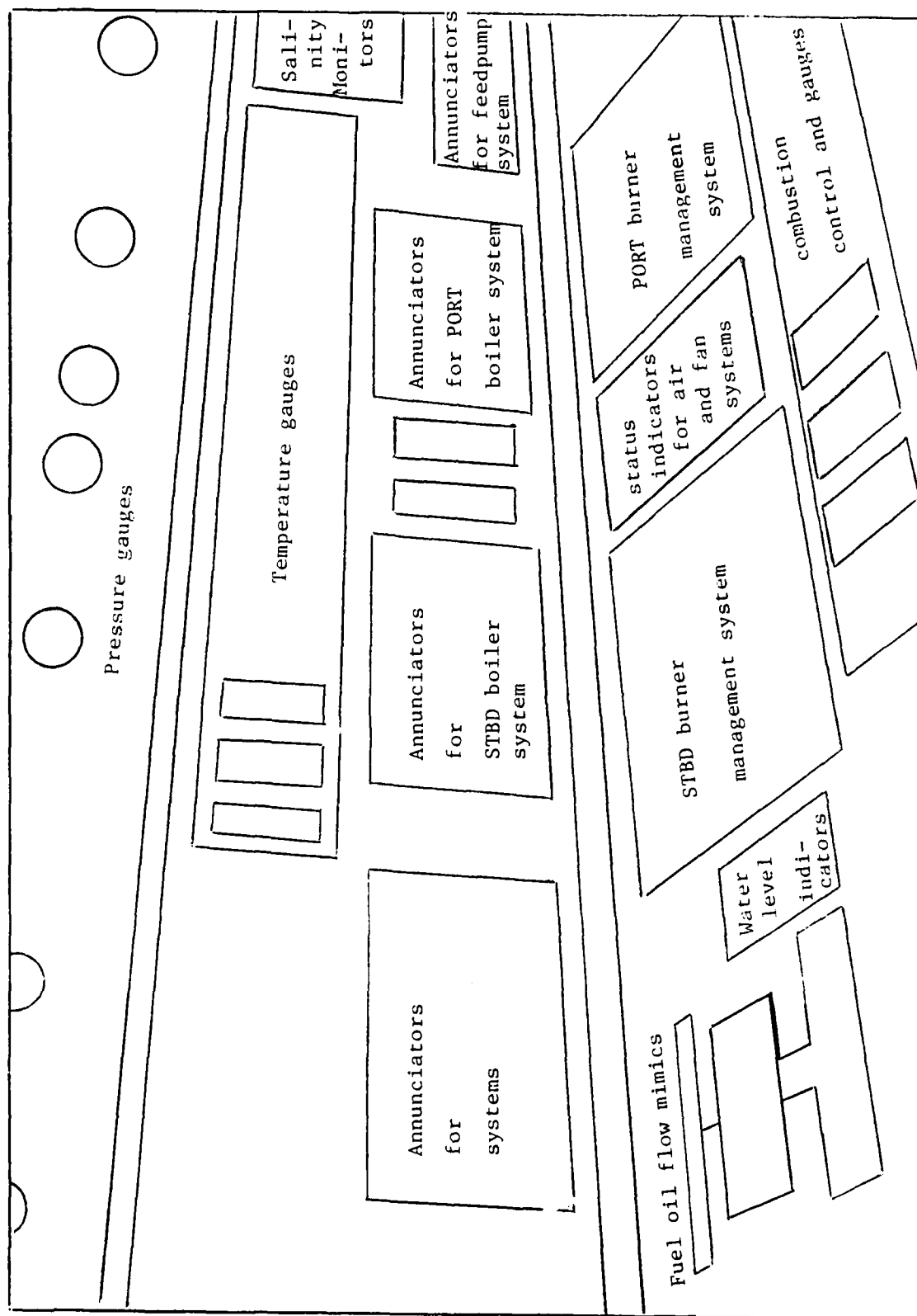


Figure 3-1 (Continued): Legends for the HFS Console.

Table 3-1: Comparison of three simulators

	Physical Fidelity	Structural Fidelity	Dynamic Fidelity
Low Fidelity Simulator	Low	Moderate	Low
Moderate Fidelity Simulator	Low	Moderate	Moderate
High Fidelity Simulator	High	High	High

The Low Fidelity Simulator -- FAIL

This low fidelity simulator runs on an APPLE II Plus microcomputer, with 64K bytes of memory, an 80-column card and a real time clock. It is used to accept inputs from the subject and to display powerplant status and other relevant information (Figure 3-2).

In addition, a number of hardcopy diagrams provide system schematics (e.g., Figure 3-3) (see Appendix D) at various levels of detail. Also provided is a list of commands (see Appendix F) used interactively to check for the status of gauges, alarms and indicators.

The simulator evolved from FAULT (Framework for Aiding the Understanding of Logical Troubleshooting) (Hunt and Rouse 1981). FAULT is an interactive context-specific troubleshooting simulator. In FAULT, the

Operation		Symptoms	
Gauges	G	Initial Condition : Full ahead at 90 RPM. Symptoms : Vacuum Low Alarm Sounds. Vacuum drops when slowing down the ship.	
Alarms	A		
Indicators	I		
Look	L		
Diagnosis	D		
Quit	Q		
Recall	R		
Gland seal steam pressure low		(II)	
Action :D168			
G35 C=3 G60 C=3 G31 C=3 I9 C=2 L77 C=10 R D168 C=50			(IV)

Figure 3-2: Main Screen Layout of FAIL.

- I: Actions available.
- II: System message display window.
- III: Action designation window.
- IV: Record of past actions and costs.
- V: Initial condition, symptoms, indicator status and recalled information.

user is given a hardcopy form of a network representation of the system of interest, and a statement of the symptoms which reflect the fact that some component or part in the system is malfunctioning. By interacting with a CRT terminal, the problem-solver must gather information in order to identify which of the components or parts has failed. Problem solving performance is judged by the total cost to identify the failed component, calculated from the costs of all actions and parts replaced. The cost for each simulated action corresponds to the actual cost (or time equivalent) of performing such an action on a real system.

Display Interface

Using design concepts similar to those used in FAULT, FAIL presents failures to the subjects via a CRT terminal. All the problems used were prepared by the MSI experts. They were real failures. Each problem presented a snapshot of a failure. At the start of a trial, initial conditions and the symptoms of a failure are shown on the display (Figure 3-2). For example, the initial condition could be "Steaming at sea", and the symptoms of failure could be "Vacuum pump fail alarm sounds". The subjects may request more information by typing in appropriate commands.

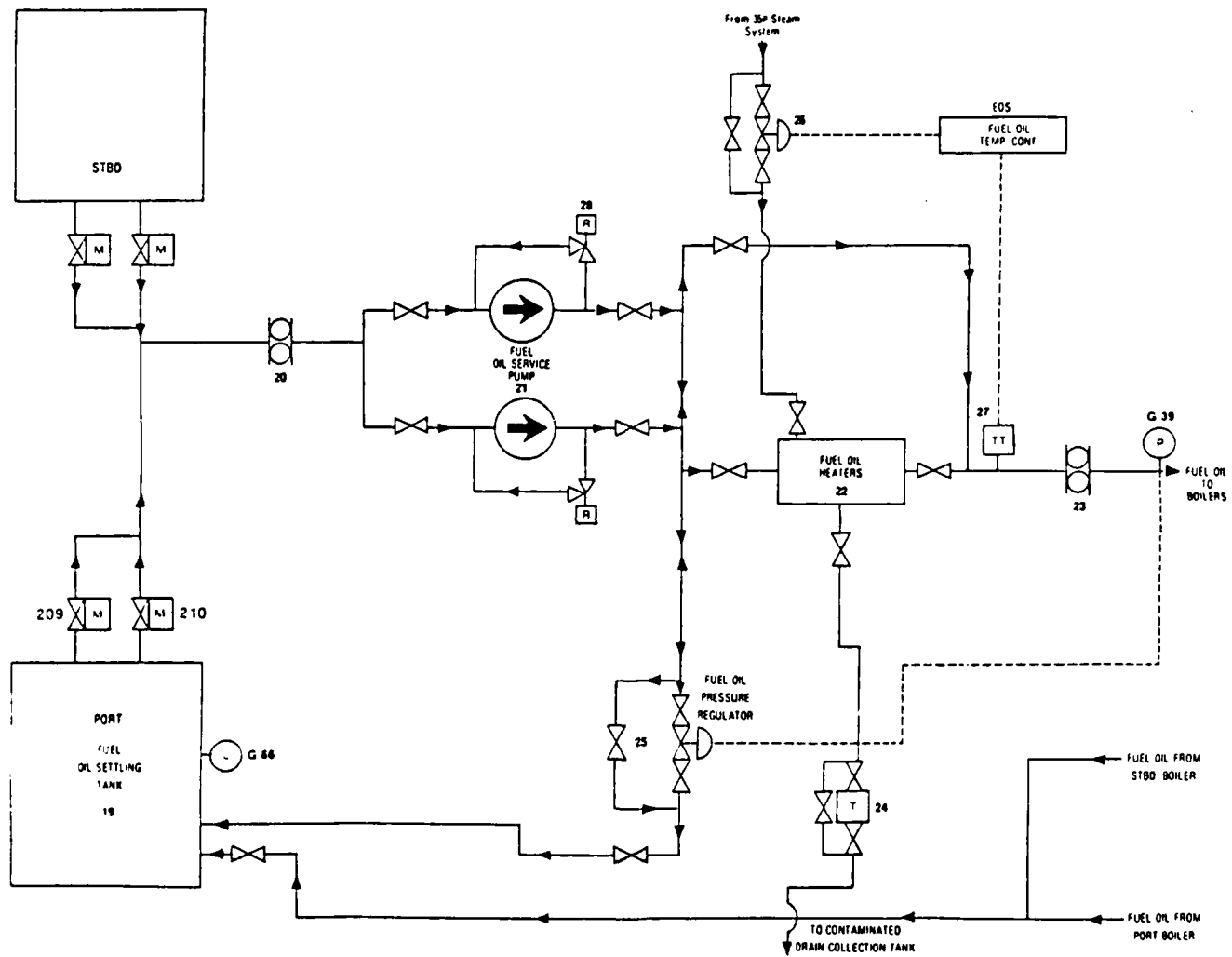


Figure 3-3: FUEL OIL SUPPLY SYSTEM

A subset of the actual alarms, status indicators and gauges (384 out of approximately 500) is used in this simulator. Gauges, alarms or status indicators for subsystems that are not power-related, such as service water pressure, were not included. Displays other than gauges and meters are classified into 12 groups by their location and function. Each group is identified by a name. Alarms and status indicators are given in groups while gauges are shown individually because approximately the same effort is required to locate the former in a group, or the latter separately.

All actions and their associated costs are displayed at the lower left portion of the screen. Cost for each action reflects the time involved to obtain the relevant information on the real system. Estimated values of various costs were provided by marine engineering experts at MSI. The objective is to minimize the cost, the actual time spent and the total number of actions for diagnosing the failure. At the end of a trial, the subject's total number of actions, cost and total time elapsed are displayed.

The subject requests information about gauges, alarms and status indicators by typing in appropriate commands. Only the alarms and the status indicators that are activated by the given failure are shown to the subject to simulate the illuminated tiles in the real system. The subject keeps requesting information from the system until he is ready to diagnose. Then he identifies the failed component by designating the number shown beside the chosen component on the schematics. Gauges are also numbered, as well as color coded on the schematics to avoid confu-

sion with component numbers.

Operator Inputs

Each of the seven possible actions is described below.

1. GAUGES (G) :

A qualitative description such as LOW, HIGH, NORMAL or ABNORMAL is shown instead of the actual readings of the gauges. When appropriate, the time history of the gauge of interest can also be provided. For example, a message like "Pressure drop to 300 lb in 10 min" could give the subject some idea about changes of the system states over time. The cost associated with this action is 3 units.

2. ALARMS (A) :

This option displays the alarm status of a particular subsystem. Only those illuminated tiles in the real system are shown in the right half of the display. If none of them is turned on, an "ALL ARE NORMAL" message will appear. When more than 22 alarms are on at one time, the last line displays: "MORE ON NEXT PAGE. PRESS "ENTER" TO SEE MORE". The subject presses either the "ENTER" key, or types "Q" to return to the command mode. The cost associated with this action is 2 units.

3. INDICATORS (I) :

The status indicators of the subsystem, that provided information such as "on", "off", "standby", or "running", are shown on request. The content and format are similar to "ALARMS". The cost is 2 units.

4. DIAGNOSIS (D) :

The users can check their conclusions and/or test their hypotheses using this option. All the failures can be traced down to the physical component level using numbers on the system schematics. The cost is 50 units.

5. LOOK (L) :

This option simulates the action of going into the engine room to physically inspect a component. The subjects can observe the status of various components including noises, gauge readings and vibration. No complicated procedures are involved to inspect a component under this option. For example, to respond to a "LOOK" at the feedwater regulator, the system may indicate "open 10 %". If the components the subjects choose to "LOOK" at cannot be seen on a real ship such as "boiler tubes", "sea scoop" etc, the system responds with "not available". The cost is 5 units.

6. QUIT (Q) :

This option is used to terminate the diagnostic process without completion. No cost is assigned to this option.

7. RECALL (R) :

Since all the actions and their responses are stored, they can be reviewed using this option. No extra cost is incurred for doing this.

An Example Run

A typical problem session with FAIL can be illustrated using Figure 3-2. Figure 3-2 shows the main screen for Failure 13 in which the gland seal regulator failed (component no. 168). The record of actions performed by a subject is shown in region IV. At the beginning, the initial condition and symptoms are shown in region V. They may be removed to make space for indicator status etc. After studying the symptoms, the subject decided to check the cooling water pressure. He searched the command list and found that Gauge 35 provided its value. Then he typed G35 in region III. The system responded by displaying "Main cooling water pressure normal" in region II.

He then checked the condenser level (G60). The message "Main condenser level normal" was displayed in region II. The subject typed in "G31" and saw "Gland seal steam pressure low" in region II. To know more about the condenser system, he chose "I9" (Indicator status of condenser system). FAIL displayed all relevant indicator status such as "pump running", "valve closed" etc., in region V. After studying the indicator status of the condenser system, he might hit the carriage return key to move the cursor back to region III for further actions.

The next command was "L77" which revealed in region II the status of a component, vacuum pump, in the engine room. The following "R" command displayed all the previous information in region V for review. He diagnosed the failed component correctly via the last command "D168". FAIL responded to the correct solution by showing statistics such as total time spent, total number of actions and total cost. This

statistics would also appear if "Q" was chosen. Any time the correct solution or the "Q" command was chosen, a problem was terminated. FAIL would then go to the next problem.

Summary

FAIL is a low fidelity simulator of a marine powerplant. It was used in experiments as an integral part of a training program at MSI. Subjects, who were participants in a training program at MSI, were presented with realistic failures on FAIL. They were asked to find the cause of the failure while keeping the total time spent and the total number of actions small. Subjects could interrogate FAIL to find out the system status, which might help the generation and evaluation of diagnostic hypotheses. All actions and their associated costs were recorded for later analysis. A sequence of experiments were conducted. The next chapter describes the design of the experiments, including a description of the subjects, the experimental procedures and the data collection details.

CHAPTER IV

DESIGN OF THE EXPERIMENTS

A low fidelity simulator, FAIL, used in experiments to study fault diagnosis behavior was described in the previous chapter. FAIL incorporates realistic failures and system schematics to simulate failure situations in a marine powerplant. An experimental session consisted of a number of problems. Each problem provided a set of symptoms caused by the failure of a single component. Subjects were asked to find the faulty component that caused the failure symptoms presented to them and at the same time minimize the real time spent and actions. The following is a description of the subjects, the experiments and the data collection procedures.

Subjects

Engineers from various fleets of a major petroleum processor who attended an MSI training course participated as subjects. They represented American, English, and Italian fleets. All were experienced marine engineers. Twenty-one out of 28 were chief engineers. The rest were either first, second or third engineers. The chief engineers had more experience than the first engineers, who in turn had more experience than the second engineers.

The training course is designed to improve their problem solving skills. It is two weeks long and consists of three parts: HFS training, laboratory training and lectures. The lectures were given during the

second week, while the subjects alternated between the HFS training and laboratory training in the first week. During the laboratory training, subjects were given exercises in logical troubleshooting computer tasks using TASK (Rouse 1981), FAULT (Hunt and Rouse 1981) and FAIL. FAULT is a context-specific troubleshooting task as described in Chapter III. TASK involves fault diagnosis of graphically displayed network of nodes. Each node may have random number of inputs and outputs. If a node fails, it will produce values of 0 on all the outputs that emanate from it. Given the output status of the last column of nodes, trainees are asked to find the faulty node. Contrasted with FAULT, TASK is a context-free troubleshooting task.

Subjects varied greatly in their experience with and exposure to computers. Most had never seen or used any type of computers before. They were given TASK and FAULT for the first two days of the first week (approximately three hours a day) to familiarize them with the use of computer keyboard, monitor and disks, as well as simulated diagnostic tasks. On the third day, a set of three failures were used in a demonstration to familiarize them with FAIL. Since they had experience on TASK and FAULT, this brief training was sufficient.

The Italian subjects had some language difficulties. They were allowed to ask questions about the problems. Help was sometimes provided by a MSI expert who speaks fluent Italian when they had language difficulties.

Experimental Procedures

The subjects were seated before APPLE II Plus microcomputers in a room that accomodates eight people comfortably. This room is used exclusively for problem sessions using APPLE II computers. They were given a list of components, a set of schematics of the system, and a command list (see Appendix F). An instruction sheet (see Appendix E) was read to them which explained the experimental procedures. For each trial they were asked to find the solution by identifying the failed component while keeping the total cost, number of actions and time to a minimum.

In order to study fault diagnosis behavior in a realistic diagnosis environment, experiments were designed such that the subjects would not be unnecessarily constrained by the study. The following procedures were adopted to achieve this goal.

1. The experiments were integrated into a training program designed to improve subjects' fault diagnosis skills. The program was two weeks long and consisted of three parts: HFS training, lectures, and laboratory training. FAIL was run as part of the laboratory training.
2. Instead of calling the experiments "experiments", we referred to them as "exercises". This produced a positive attitude towards the experiments. However, some subjects got too serious and asked if the results of the "exercises" would be sent back to the company for evaluation. To reduce their anxiety, they were told that these

"exercises" were designed solely for this training program, and nothing would be sent back to the company.

3. Subjects were asked not to discuss the "exercises" with each other. They were told that all their questions would be answered by the experts at MSI and discussed openly in the classroom at the end of the training program.

Experiments and Data Collection

A pilot experiment and five experimental sessions were conducted. From the pilot experiment, it was found that the command "A" for checking the alarm status was not useful since most of the salient symptoms had already been presented on the initial screen. On the other hand, we found that an option to simulate the behavior of checking the physical status of components in the engine room could increase the realism of FAIL. A new command "L" (for "Look") was then designed and incorporated into FAIL for this purpose. Therefore, from Session 1 on, command "L" took the place of command "A". It was also found from the pilot experiment that some of the failures (see Appendix A) were not realistic. A different set of failure (see Appendix B) was compiled by experts at MSI from reports of real failures. Throughout the experiments, the actions taken by subjects were recorded automatically by FAIL.

No models or frameworks of fault diagnosis were assumed before the design of FAIL and the design of the experiments. A preliminary description of fault diagnosis was to be based on observations of subject performance. A framework of fault diagnosis performance was

developed using the data collected and observations made from experiments. The framework was then used to identify and verify factors affecting fault diagnosis performance from Sessions 1 through 5. This analysis quantified the quality of actions and then averaged these data across subjects.

Individual variation is such a striking feature of human fault diagnosis performance that any attempt to average data across a population is certain to mask features of the experts' behavior. It is essential to get enough data about each individual subject to identify what information he has and how he processes it. As Newell and Simon (1972) point out, only the verbal behavior can dig deeply into the complexity of cognitive behavior. Therefore, for Sessions 4 and 5, in addition to the regular FAIL diagnostic task, some subjects were asked to "think aloud" and the protocols were recorded. Kuipers and Kassirer (1984) argued that this type of think-aloud experiment is particularly sensitive to the natural control structure of the subjects' problem solving method. The information reported was actually in the subjects' focus of attention at the time of diagnosis. A model of fault diagnosis is proposed based on these protocols and data from experiments. Details of this model are discussed in Chapter VI.

Descriptions of the experiments are given below.

Pilot Experiment

Five subjects participated in this experiment; They are referred to as S1, S2, S3, S4 and S5 later in this thesis. The failures used on

HFS were different from those on FAIL. Each subject was given 18 problems (see Appendix A) containing one failure each. The first three problems were used for training. Problems were presented to each subject in the same sequence. Subjects were allowed to spend as much time as they wanted on each problem. Trials were sometimes interrupted because of the HFS schedule. The sequence was continued later at the point where it was interrupted. Subjects were allowed to ask questions since some of them had language difficulties. The "L" command was not available for this experiment.

Session 1

Four subjects participated in this session (S6, S7, S8, S9). This session used a set of 29 failures (see Appendix B), which were different from those used in the pilot experiment. By repeating five of the failures, a total of 34 problems were presented. The sequence of problems varied for each subject. The first three problems were used as training trials. The "L" command was available in this session.

Session 2

Five subjects participated in this session (S11, S12, S13, S14, S15). They were less experienced than trainees in the pilot experiment and Session 1. Most of them were either second or third engineers while the trainees in the pilot experiment and Session 1 were chief engineers. The set of failures used in Session 1 was used in this session except that only Failure 1 and Failure 3 were used for training and Failure 9 was modified slightly. The experimental environment was very similar to

that of Session 1. However, the sequence of problems given to each subject was the same for all subjects.

Session 3

Seven subjects participated in this session (S16 - S22). All were chief engineers. The set of failures used in Session 1 was used in this session.

Session 4: protocols taken

Six subjects participated in this session (S25 - S30). All were chief engineers. The set of failures used in Session 1 was used in this session.

Session 5: protocols taken

Six subjects participated in this session (S31 - S36). They were either chief, first or second engineers. The set of failures used in Session 1 was used. However, Failures 27 and 29 were used for training instead of the first three. Also, the sequence of failures used in Sessions 2, 3, and 4 was reversed.

Summary

Twenty-eight expert engineers participated in five experimental sessions as subjects. Several modifications were made based on the pilot experiment, including commands available and set of failures used. Precautions were taken to ensure that subjects had positive attitudes

towards the experiments. Actions were recorded automatically for analysis which quantified the quality of actions and averaged across subjects. To investigate diagnosis behavior more deeply, some subjects in Sessions 4 and 5 were asked to think aloud while solving problems. The experimental results were analyzed to probe problem solving behavior. The next chapter describes a framework for fault diagnosis. This framework is used to analyze the data from Sessions 1 through 5. The implications of these experimental results are also discussed.

CHAPTER V

ANALYSIS OF EXPERIMENTAL RESULTS

In the preceding chapters we described a low fidelity simulator of a marine powerplant, FAIL, and how it was incorporated in MSI training program to run experiments. Data collected from the last five experimental sessions were analyzed with help from subject-matter experts at MSI. This chapter describes the results of the data analysis. First a framework for fault diagnosis is described. Using this framework, two factors affecting the fault diagnosis performance are identified and tested for statistical significance. The implications of these factors on training are discussed. Then, some observations on subjects' responses are described.

A Framework for Fault Diagnosis

Fault diagnosis can be viewed as a backward reasoning process that goes from symptoms to causes. However, the knowledge accumulated from years of education is arranged in a forward reasoning format that goes from components (hence causes) to symptoms. Subjects are well versed in reasoning from a failed component to its symptoms based on the system topology and dynamics. Given sufficient information about the input conditions of a component, expected output states or influence on the downstream components can be derived if appropriate dynamics is used (see Figure 5-1). Knowledge of system dynamics is the basis of this forward reasoning process. Though the dynamics may be complicated, the reasoning from causes to symptoms is relatively straightforward.

However, backward reasoning from symptoms to components cannot be accomplished by applying the system dynamics directly because rules of dynamics do not work in reverse time order. In this backward reasoning process, the counterpart of system dynamics, and hence the output (the cause), is unknown (see Figure 5-2).

When a failure occurs, the faulty component causes deviations from normal system states. Among these state deviations some are obvi-

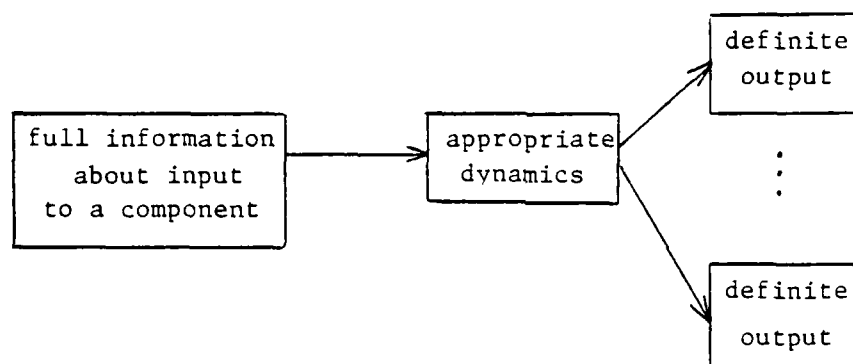


Figure 5-1: Forward Reasoning

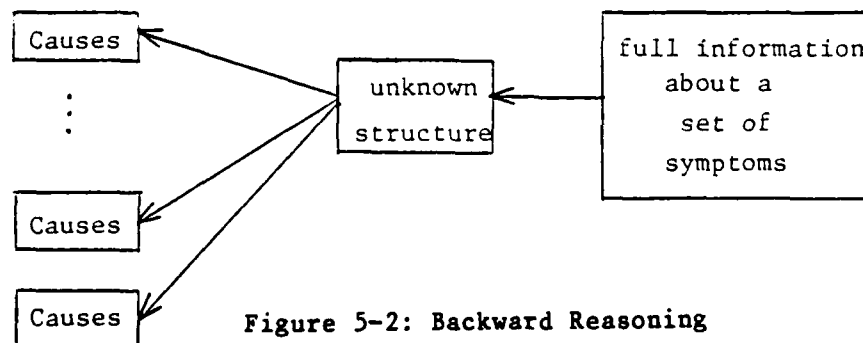


Figure 5-2: Backward Reasoning

ous and manifest themselves in the form of alarms, or other irresistible clues, while some are non-obvious and must be investigated. We call the first kind the "obvious" symptoms and the second kind the "non-obvious" symptoms. In most cases, the obvious symptoms are just alarms representing the non-obvious symptoms that exceed acceptable levels.

Obvious Symptoms, Non-obvious Symptoms, and Causes

Obvious symptoms are those presented usually via audio alarms or flashing lights to attract the operator's attention. Compensatory actions appropriate to these symptoms are often required to maintain the powerplant in an operating mode. However, these obvious symptoms alone may not provide enough information to diagnose the failures. The subjects need to gather more information in order to form a hypothesis about the faulty component and evaluate the hypothesis. To gather more information the subjects must interrogate the system for the non-obvious symptoms.

The relationship between symptoms and causes is dynamic since the symptoms could change with time. However, at any moment, a rather simple relationship between symptoms and causes seems to exist. That is, a set of obvious symptoms may be common to several sets of non-obvious symptoms each resulting from several different causes (Figure 5-3). A set of causes is the collection of components which are suspected to be the cause of observed symptoms. It is possible that several different causes may lead to the same set of non-obvious symptoms. A set of causes may take the form of a subsystem that is characterized by some clearly identifiable function. For example, it may be the "Condenser

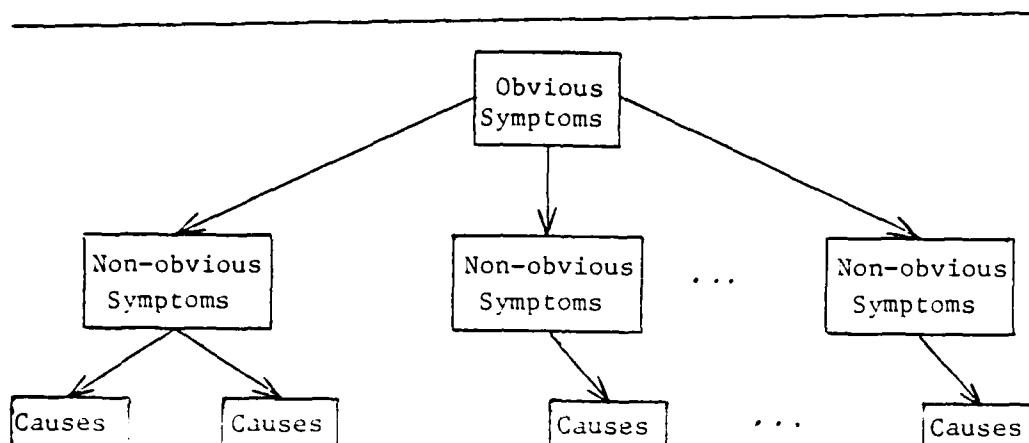


Figure 5-3: A Symptom-Cause Tree

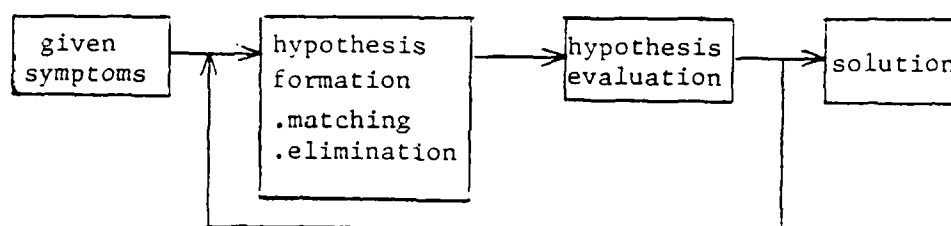


Figure 5-4: Stages of Diagnostic Process

System" that removes heat from steam to reuse water. The "Condenser System" consists of several subsystems such as condensate system, circulating water system, condenser etc. Each subsystem can be further decomposed into smaller subsystems or components.

Hypothesis Formation, Hypothesis Evaluation and Stage Shifting

The apparent relationship between symptoms and causes actually resembles an inverted tree with obvious symptoms at the root and the

causes forming the leaves. The diagnostic process is to find a path from the root to the correct leaf, i.e., the faulty component. This process can be further broken down into two stages. In the first stage, given obvious symptoms the subjects gather information about the system in order to find a set of possible causes that explain all the known symptoms. These symptoms include relevant information uncovered during the investigation. The available information might result in the elimination of certain candidates from the suspected set of feasible causes. This is like forming branches for the symptom-cause tree. After a set of possible causes is identified, the subjects shift into the second stage in which they try to identify a narrower set of causes and search through the hierarchy of the chosen set of causes to the faulty component. Results from this stage may be fed back into the hypothesis formation stage to eliminate infeasible candidates. In other words, in the first stage, hypotheses about faulty subsystems are formed while in the second stage, these hypotheses are evaluated and tested. Subjects alternate between these two stages to arrive at the solution (Figure 5-4).

Feasible Sets

A preliminary step before any hypothesis can be formed is to search through the knowledge base to find related symptom-cause pairs. This is done to build new branches for the symptom-cause tree. The collection of these branches forms a feasible set of causes for the given failure. When members of this feasible set related to the failure are found, new hypotheses are formed. The feasible set is modified as the

diagnostic process continues. Older branches are discarded if evidence gathered does not warrant further consideration. New branches are added as more elaborate reasoning is completed. Although each hypothesis may appear reasonable, not every hypothesis is guaranteed to be correct. The failed hypothesis may provide insights into localizing the failure, and for forming new hypotheses.

Factors Affecting Fault Diagnosis Performance

Two factors affect the efficiency of a diagnostic process. They are the initial feasible set and the strategies that govern the transition from the hypothesis formation stage to hypothesis evaluation stage. The identification of the initial feasible set, the transition strategies and the analysis of the experimental results are discussed next.

Initial Feasible Set

The initial feasible set (IFS) consists of the symptom-cause pairs that the subjects considers before any hypothesis is formed and evaluated. It reflects the strongest symptom-cause pairs that are associated with the given obvious symptoms.

To study the relationship between the initial feasible set and overall performance, an operational definition and measure for initial feasible sets were necessary. It was impractical to ask the subjects what their initial feasible sets were. An operational criterion was to take the first three actions as the initial feasible set if no stage shifting occurred. Stage shifting was assumed to have occurred if a subsystem or a component was diagnosed as the cause of failure. If

subjects transitioned to the hypothesis evaluation stage before their third actions, only the actions before transition were taken as the initial feasible set. According to this criterion, members of an initial feasible set might be a component check, gauge reading or alarm status inquiry.

If the initial feasible set contained components that were related to the faulty subsystem or component, the set was rated as "good", otherwise "bad". This description of "good" and "bad" is used as a convenient gross measure without implying any measure of the amount of "goodness".

Table 5-1 shows examples of initial feasible sets that were formed by subjects for Failure 1-6. Since the faulty component was the gland seal regulator, the initial feasible sets formed by S1, S4 and S5 were rated as "good" and those by S2, S3 were rated as "bad".

Transition Strategies

Two types of strategy were observed to result in transitions from the hypothesis formation stage to the hypothesis evaluation stage. They were the breadth-depth strategy and the balanced strategy.

(1). Breadth-Depth Strategy (BD):

Subjects stayed in the hypothesis formation stage until further symptoms were found and hypotheses were formed. Then they switched into the hypothesis evaluation stage. They thoroughly tested these hypotheses before they gave up and went back to the hypothesis formation stage. In other words, they conducted a broad search for

Table 5-1: Initial Feasible Sets for Failure No. 1-6

Failure 1-6: Low vacuum alarm sounds, vacuum steadily dropping on slowing to half ahead during maneuvering.

Faulty component : Gland seal regulator

Subject	Initial Feasible Set	Rating
S1	gland seal steam pressure main cooling water pressure	good
<hr/>		
	150 psig aux system pressure	
S2	condensate pump pressure main cooling water	bad
<hr/>		
S3	main condenser level condensate pump pressure	bad
<hr/>		
S4	gland seal steam pressure	good
<hr/>		
S5	gland seal steam pressure	good
<hr/>		

candidates and did thorough checking once a hypothesis was formed.

(2). Balanced Strategy (BL):

Subjects did not conduct a broad search for candidates in the hypothesis formation stage. They switched to the hypothesis evaluation stage although no obvious hypothesis was formed. They did not seem to evaluate any single hypothesis thoroughly. They tended to form hypotheses quickly and maintain several hypotheses at the same time.

To analyze the data, two difficulties had to be resolved.

(1). how to distinguish the hypothesis formation stage from hypothesis evaluation stage?

(2). how to identify the dominant strategy?

The identification of stages was very difficult. The following criteria were used:

(1). If more than two actions for a particular subsystem were checked consecutively, then it was assumed that the subjects were in the hypothesis evaluation stage.

(2). If a hypothesis had been tested before, or a particular component or gauge had been considered before, then later related investigations were considered as entering the hypothesis evaluation stage.

(3). Actions that could not be classified using the two criteria above were regarded as actions in the hypothesis formation stage.

Two more criteria were used to help classify the strategy as balanced or

breadth-depth.

- (4). If multiple hypotheses were observed, then more weight should be placed on balanced strategy.
- (5). If each hypothesis evaluation stage consisted of more than three related actions, then more weight should be placed on breadth-depth strategy.

The following examples illustrate how these criteria were used to identify stages and hypotheses. An asterisk before an action denotes that it is fault-related.

Example 1: Bad IFS and BL

Failure 1-13 : condenser vacuum low alarm sounds

Subject : S1

Actions recorded :

1. G60 Main condenser level (IFS)
2. G36 Condensate pump pressure (IFS)
3. G34 Salt water service pressure (IFS)
4. * G27 Condenser vacuum pressure (fault related action)
5. * G35 Main cooling water pressure (fault related action)
6. D83 Sea scoop valve (hypothesis 1: circulating water)
7. D78 Main condenser (hypothesis 2: condenser)
8. D84 Sea scoop (hypothesis 1)
9. D80 Main circulator discharge valves (hypothesis 1)
10. G35 Main cooling water pressure
11. G19 Deaerator tank pressure (hypothesis formation)

- 12. G55 LP turbine exhaust temp (hypothesis formation)
- 13. G15 PORT FO to burners pressure (hypothesis formation)
- 14. I5 Salinity system (hypothesis formation)
- 15. G80 No. 1 Deaerating feedtank dump regulator
(hypothesis 3: deaerator tank)
- 16. D82 Main circulator suction valve (hypothesis 1)
- 17. D87 Sea strainer (hypothesis 1)

Total number of actions: 17

Total cost: 323

In example 1, actions 1, 2 and 3 were classified as the initial feasible set. Since none of them were related to the cause of the failure, this IFS was rated as "bad". Actions 4 and 5 were additions to the feasible set. These two actions were related to the fault directly. Action 6 showed the first hypothesis which was related to circulating water. Action 7 showed another hypothesis which was concerned with the condenser itself. Actions 8 and 9 tested the first hypothesis further. From action 10 to action 14 the subjects switched back to the hypothesis formation stage. Action 15 was related to action 11 and was thus regarded as the third hypothesis which dealt with the deaerating tank. In actions 16 and 17, the subject picked up the first hypothesis again and obtained the solution. The subject maintained three hypotheses at the same time. He alternated among hypotheses. Therefore, it was classified as using a balanced strategy.

Example 2 : Good IFS and BD

Failure 1-13 : condenser vacuum low alarm sounds

Subject : S2

Actions recorded :

1. * G35 Main cooling water pressure (IFS)
2. D81 Main circulators (hypothesis 1: cooling water)
3. R0 Recall
4. D83 Sea scoop valve (hypothesis 1)
5. D85 Sea high suction (hypothesis 1)
6. D86 Sea low suction (hypothesis 1)
7. D87 Sea strainer (hypothesis 1)

Total number of actions: 7

Total cost: 253

In example 2, the initial feasible set contained only one member which was related to the cause of the failure. Therefore the IFS was rated as "good". From action 2 onwards, the subject formed a hypothesis about the cooling water system and conducted thorough checking on it until action 7 when the solution was found. The subject maintained only one hypothesis and did thorough checking on that hypothesis. It was thus classified as using a breadth-depth strategy.

Analysis of Experimental Results

Of 29 failures, 21 were considered for detailed analysis. Among the failures not considered, five were used for training (1, 2, 3, 27, and 29), two resulted in confusion due to ambiguous symptoms (18, 23), and one was merely a duplicate (28) of another failure (13).

Twenty-eight marine engineers participated as subjects in five

experimental sessions. Since the experimental conditions were very similar, and the same set of failures was used throughout all experiments, data were analyzed across the subjects from different experiments.

The analyses were carried out along four different directions. First, the failures were classified into different difficulty levels. Second, performance between subjects using good IFS and bad IFS was compared. Third, performance between subjects using BD strategy and BL strategy was compared. Finally, performance from subjects using both good IFS and BD strategy was compared with the rest. These are described below in detail.

Easy vs. difficult failures

Means and standard deviations of number of actions and time were computed for each failure. The failures are plotted using means of the total number of actions and time in Figure 5-5. Three distinct clusters can be seen. Failures that required more than 700 seconds were classified as difficult. Of the remaining failures, those requiring more than 10 actions were classified as moderately difficult; the rest were classified as easy. Cluster 1 contains Failures 7, 8, 9, 10, and 19. Cluster 2 consists of Failures 16, 20, 21 and 26. The remaining failures, e.g., Failures 4, 5, 6, 11, 12, 13, 14, 15, 17, 22, 24, and 25 are in cluster 3.

The easy failures were not considered for further analyses because the diagnostic processes for easy failures were too short to be

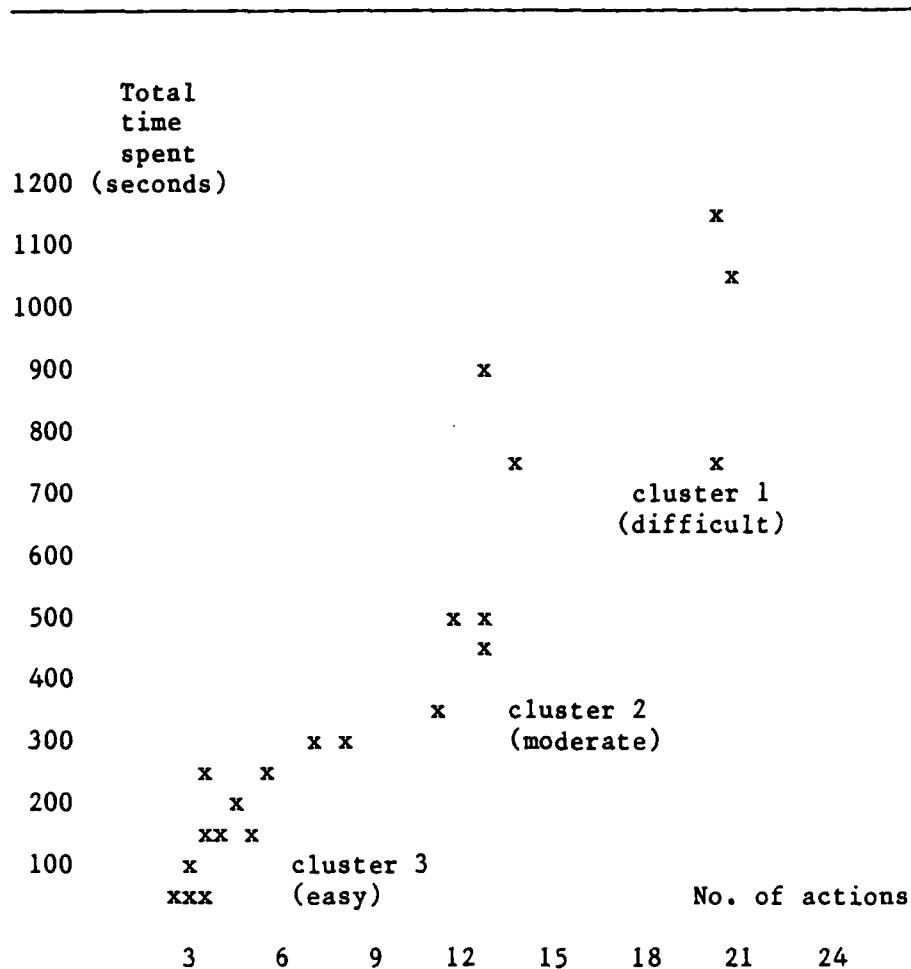


Figure 5-5: Performance based on total time spent and total number of actions.

Cluster 1: 7, 8, 9, 10, 19
 2: 16, 20, 21, 26
 3: 4, 5, 6, 11, 12, 13, 14
 15, 17, 22, 24, 25

informative.

Table 5-2 lists the statistics for each failure under both measures. Two observations are relevant: large standard deviations and varying sample sizes. The varying sample size reflects failures where

Table 5-2: Performance statistics.

Failure no.	Diffi- culty	Sample size	No. of Actions		Total Time (sec)	
			mean	sd	mean	sd
4	Easy	28	6.0	3.1	225	178
5		28	4.3	2.6	167	126
6		27	1.9	1.5	82	73
11		25	5.8	6.2	182	173
12		21	9.0	8.2	297	272
13		27	6.6	7.5	295	416
14		25	4.6	3.6	163	102
15		26	5.2	3.8	195	128
17		26	4.3	5.2	215	303
22		27	2.7	3.5	101	101
24		28	3.2	2.3	84	49
25		28	2.5	1.2	91	56
16	Mode- rate	19	12.5	9.5	445	356
20		23	12.2	12.1	499	464
21		26	10.7	8.5	318	273
26		26	11.4	7.1	484	484
7	Diffi- cult	25	21.3	14.6	1069	641
8		19	12.7	11.0	893	1037
9		20	19.5	14.7	748	780
10		17	14.8	13.6	753	1001
19		17	19.9	15.5	1106	1296

the subjects used the "Quit" command to discontinue the diagnostic process without finding the failed component. Since data from such trials were not used for analysis, the sample sizes were reduced. Difficult failures had smaller sample sizes than moderately difficult or easy failures.

Large standard deviations in performance measures could be the result of computing overall statistics without regard to the IFS and/or stage shifting strategies. The influence of these two factors is investigated next.

Good IFS vs. bad IFS

Subjects were divided into two groups: one with good IFS, the other with bad IFS. Statistics were computed for each failure from clusters 1 and 2 based on the raw data. In addition, a ranked t-test was also done based on a paper by Conover and Iman (1976). Table 5-3 summarizes the results.

It is clear that subjects with good IFS performed better than with bad IFS. 13 out of the 18 t-tests taken from raw data are significant at 0.05 level. The statistics from ranked data are even better, 17 out of the 18 t-tests are significant at 0.05 level.

BD strategy vs. BL strategy

Instead of classifying subjects by the IFS, they were divided into two groups according to the strategies used. Statistics from the raw data as well as ranked data were computed and t-tests were done

between BD strategy group and BL strategy group. Table 5-4 summarizes the results. Results were similar to what was observed above.

Subjects using BD strategy performed better than those using BL strategy. 17 out of 18 t-tests from raw data are significant at 0.05 level, while all t-tests from ranked data are significant at 0.05 level.

Good IFS and BD strategy vs. all the others

Subjects were divided into two groups for testing: one with both good IFS and BD strategy, the other the remaining subjects. The data were analyzed as before. Table 5-5 summarizes the results. It is seen that the group with good IFS and BD strategy performed better than the others.

Table 5-3: Performance comparison between good IFS and bad IFS. A: number of actions. T: Time. "*": designates a t-test for the means that was not significant at 0.05 level.

Failure no.	data type		good IFS			bad IFS			t-test	
			mean	sd	N	mean	sd	N	t-score	v
16	raw	A	7	3.9	10	18.6	10.4	9	3.16	10
		T	265	212		646	385		2.63	13
	ranked	A	13.3	4.2		5.3	3.5		4.50	19
		T	12.5	4.7		7.1	5.3		2.33	18
20	raw	A	9.6	11.3	18	21.8	10.6	5	2.25	8
		T	398	453		862	320		2.60	11
	ranked	A	12.3	5.4		6.4	3.9		2.76	11
		T	13.6	6.7		6.2	3.4		3.35	17
21	raw	A	7.9	6.1	21	22.2	7.8	5	3.82	6
		T	243	133		637	472		1.85*	4
	ranked	A	15.1	6.4		4.4	3.5		5.10	14
		T	15.1	7.0		6.2	6.9		2.60	7
26	raw	A	8.2	4.8	13	14.7	7.7	13	2.59	21
		T	472	500		495	487		.12*	26
	ranked	A	16.6	7.1		9.3	6.1		2.77	25
		T	14.5	8.5		12.5	6.9		.66*	25
7	raw	A	11.7	17.5	6	24.4	12.6	19	1.65*	7
		T	613	583		1213	603		2.18	10
	ranked	A	19.2	8.6		10.8	5.7		2.25	7
		T	18.3	6.7		11.3	6.9		2.21	10
8	raw	A	5.1	3.2	9	19.5	11.1	10	3.93	11
		T	285	231		1441	1183		3.03	10
	ranked	A	14.1	3.6		5.9	3.5		5.03	19
		T	13.4	4.0		6.8	4.9		3.22	19
9	raw	A	12.8	10.0	13	31.9	14.4	7	3.11	10
		T	429	343		1341	1029		2.28	7
	ranked	A	13	5.1		5.4	3.5		3.90	19
		T	13.2	5.1		5.4	3.6		3.98	19
10	raw	A	12.1	10.6	15	35	21.2	2	1.50*	1
		T	538	632		2370	2125		1.21*	1
	ranked	A	9.6	4.7		3	2.8		2.82	3
		T	9.8	4.8		3	2.8		2.90	4
19	raw	A	3.3	.6	3	23.4	14.8	14	5.07	13
		T	77	15		1326	1331		3.51	13
	ranked	A	15.7	.6		7.4	4.0		7.43	15
		T	16	1		7.5	4.2		6.76	17

Table 5-4: Performance comparison between BD strategy and BL strategy.
 A: number of actions. T: Time. "*": designates a t-test
 for the means that was not significant at 0.05 level.

Failure no.	data type		BD strategy			BL strategy			t-test	
			mean	sd	N	mean	sd	N	t-score	v
16	raw	A	7.3	3.3	12	21.4	10.1	7	3.59	7
		T	251	214		779	301		4.09	11
	ranked	A	13	3.7		3.6	1.6		7.67	18
		T	12.9	4.7		4.9	2.4		4.96	19
20	raw	A	5.5	9.3	14	22.7	7.7	9	4.81	22
		T	247	371		891	294		4.62	22
	ranked	A	14.4	4.2		5.8	2.9		5.88	23
		T	15.6	5.9		6.2	3.2		4.98	22
21	raw	A	7.6	4.8	22	27.3	2.2	4	12.89	13
		T	242	130		741	473		2.10*	3
	ranked	A	15.0	6.2		2.5	1.3		8.52	26
		T	15.2	6.9		3.5	2.7		5.91	17
26	raw	A	8.2	4.7	15	15.9	7.6	11	3.00	17
		T	237	109		821	593		3.22	11
	ranked	A	16.4	6.9		8.2	5.6		3.35	26
		T	17.4	5.7		8.1	6.8		3.71	21
7	raw	A	13.9	8.0	18	40.3	9.5	7	6.51	11
		T	789	517		1787	230		6.67	25
	ranked	A	16.1	5.5		4.1	2.4		7.57	25
		T	16.1	6.3		5.14	2.4		6.30	25
8	raw	A	8.1	7.1	14	25.6	9.8	5	3.67	6
		T	384	364		2320	976		4.33	5
	ranked	A	12	4.4		3.6	2.4		5.28	16
		T	12.3	4.4		3.4	2.3		5.72	17
9	raw	A	14.1	9.5	16	41	11.9	4	4.21	5
		T	471	324		1857	1129		2.43	3
	ranked	A	12.2	5.0		3	1.9		5.97	18
		T	12.4	5.0		3	2.2		5.70	16
10	raw	A	9.3	8.0	13	32.5	13.2	4	3.32	4
		T	317	242		2172	1265		2.92	3
	ranked	A	10.6	4.2		3	1.8		5.18	16
		T	10.9	4.0		2.8	1.7		5.81	16
19	raw	A	12.6	10.5	11	33.3	14.5	6	3.09	9
		T	409	398		2383	1425		3.32	6
	ranked	A	11.3	4.0		4.3	2.6		4.35	17
		T	11.7	3.8		4	2.6		4.96	16

Table 5-5: Performance comparison between good IFS and BD and all others. A: number of actions. T: Time. "*": designates a t-test for the means that was not significant at 0.05 level.

Failure no.	data type		BD & good IFS			the rest			t-test	
			mean	sd	N	mean	sd	N	t-score	v
16	raw	A	6.1	2.8	9	18.2	9.8	10	3.72	11
		T	210	129		658	365		3.64	12
	ranked	A	14.2	3.3		5.3	3.3		5.90	19
		T	13.6	3.5		6.7	5.2		3.41	18
20	raw	A	2.8	2.9	12	22.6	9.3	11	6.76	12
		T	128	139		904	327		7.30	14
	ranked	A	15.7	2.0		6	3.5		8.09	17
		T	17.3	3.9		6.1	3.5		7.29	23
21	raw	A	7	4.6	20	22.8	7.2	6	5.10	7
		T	236	133		592	436		1.97	5
	ranked	A	15.7	5.9		4.2	3.2		6.23	19
		T	15.5	6.9		6.3	6.2		3.11	11
26	raw	A	4.2	1.9	6	13.6	6.7	20	5.60	26
		T	183	105		574	517		3.17	24
	ranked	A	22.8	2.7		10.0	5.6		7.69	22
		T	20.2	5.3		11.5	7.1		3.25	13
7	raw	A	4.6	3.0	5	25.5	13.3	20	6.44	25
		T	403	308		1235	595		4.35	16
	ranked	A	22.6	2.0		10.3	5.9		7.83	24
		T	20.6	4.3		11.1	6.8		3.90	12
8	raw	A	4.3	2.0	8	18.8	10.8	11	4.39	11
		T	227	161		1378	1142		3.50	11
	ranked	A	15	2.6		6	3.3		6.60	19
		T	14.3	3.5		6.8	4.7		3.98	19
9	raw	A	12.9	10.0	13	31.9	14.4	7	3.11	10
		T	429	343		1341	1029		2.28	7
	ranked	A	13	5.1		5.4	3.5		3.90	19
		T	13.2	5.1		5.4	3.6		3.98	19
10	raw	A	9.3	8.0	13	32.5	13.2	4	3.32	4
		T	317	242		2172	1265		2.92	3
	ranked	A	10.6	4.2		3	1.8		5.18	16
		T	11.0	4.0		2.8	1.7		5.81	16
19	raw	A	3.3	.6	3	23.4	14.8	14	5.07	13
		T	77	15		1326	1331		3.51	13
	ranked	A	15.7	.6		7.4	4.0		7.43	15
		T	16	1		7.5	4.2		6.76	17

Discussion

Based on the experimental results, it appears that the initial feasible set and the transition strategies greatly affect the diagnostic performance on a low fidelity simulator of a marine engine control room. The initial feasible set reflects the knowledge base of the subjects. The better the initial feasible sets are, the better is the performance. Although it is not clear how the initial feasible sets are formed, it is suspected that this may be affected by the knowledge of the system and the structure of the internal (mental) model linking the symptom-cause pairs.

However, a good initial feasible set in itself does not seem sufficient to explain the observed performance. Good problem solving strategy also plays an important role. It was found that a breadth-depth strategy was more efficient than a balanced strategy. It is worth noting that strategies are relatively context-free, while formation of the initial feasible set is context-specific.

The fact that both initial feasible set and the strategy used affected the performance is very significant. It seems to imply that in a highly specialized field, problem solving ability is affected by both context-free and context-specific factors. This agrees with the results and models of Rouse (1983), and Rouse and Hunt (1984).

Summary and Conclusions

Several interesting observations were made during and after the FAIL experiments at MSI. The marine engineers who participated as subjects were highly enthusiastic about their interaction with FAIL. Based on their comments, and observations during the experiments it is apparent that FAIL has two major benefits as a training tool for a complex system. They are:

(1) FAIL associates causes with symptoms

In real life, the pairing of causes and symptoms is a matter of experience. FAIL condensed these experiences into an interactive computer course which is easy to use and implement. This type of association between causes and symptoms refreshes their memory about failures and sometimes updates their knowledge about the system.

(2) FAIL forces them to think

Unfamiliar failures lead them into reasoning process which they do not usually undergo in their day-to-day operations. The reasoning process forces them to reorganize or restructure their knowledge about the system.

It would be desirable to explore these issues via rigorous transfer of training experiments (Maddox, Johnson and Frey 1985). However, no plans for such experiments exist at this time due to the difficulties involved in evaluating the trainees' real world job performance on a long term basis. Similar difficulties have been reported by Adams

(1979) and Blaiwes, Puig and Regan (1973).

The feedback provided by the trainees corroborates the experimental findings to a great degree. The strong relationship between IFS and overall performance emphasizes the important role of system knowledge on diagnosis. Frequent associations between symptoms and causes may result in the trainees' developing better IFS. Better IFS should lead to better performance.

The reason why IFS is highly correlated with overall performance can perhaps be explained by the anchoring effect (Tversky 1974). Humans have a tendency or bias toward the initial thought which may create "cognitive tunnel vision" (Sheridan 1981). Therefore, the better the IFS is, the better is the performance.

The implications of these results are twofold. First, diagnostic training should emphasize the pairing of symptoms and causes. Second, it should try to discourage trainees from making wrong hypotheses in the early stages of a diagnostic task. The first is somewhat obvious and has been carried out in most training programs. However, it is not obvious how the latter can be incorporated into a training program and further investigation is needed.

Although trainees were not aware of the existence of stage-shifting, they felt that more thinking and planning could help improve the diagnostic process. Subjects who used BL strategy had a tendency to generate a large set of plausible hypotheses about failed components that could have produced the observed symptoms. However, humans are

observed to have difficulties in generating all plausible hypotheses in diagnosing (Mehle 1982). Also, due to short term memory limitations (Miller 1956), and other cognitive limitations in keeping track of more than one line of reasoning at any time, it is extremely difficult to hold several active hypotheses simultaneously. Maintaining multiple hypotheses generates confusion and heavy cognitive load, which reduces efficiency.

The relationship observed between transition strategy and overall performance indicates that context-free problem solving training is very useful. In fact, some trainees reported that they benefitted from TASK (a context-free logical troubleshooting environment) and FAULT. TASK and FAULT were administered along with FAIL to improve skills in logical inference. The implication of this for technical training is that a training program that encourages more thinking and planning rather than disorganized diagnostic behavior should be designed if at all possible.

Two important observations can be made concerning future research. The first is how dynamics of a system affects the diagnostic process. An important difference between a static system and a dynamic system is that the latter can get worse as time goes by since inappropriate actions might be taken by the subject. The framework might need to be modified to account for this phenomenon.

The second concerns individual differences among the experts. Although the trainees have very similar training on this particular system, they responded differently when presented with failures. It is conjectured that the way their experiences are organized and integrated

with system knowledge might be responsible for the differences. In other words, the process in which the knowledge is organized might play an important role in how this knowledge is used. This involves the acquisition, structuring, and retrieval of knowledge. These three aspects are highly interdependent (Baddeley 1976) and, hence, none of them can be considered individually to the exclusion of the others. A comprehensive model could possibly be developed based on research on knowledge acquisition and knowledge application discussed below. These issues and those discussed below are considered in greater detail in chapter VI.

Knowledge acquisition is concerned with how experiences help build up the knowledge base. It emphasizes the learning process and the structuring of knowledge in memory. The dynamic memory model of Schank (1982) and the long term memory model of Kolodner (1983) might prove useful to study this problem.

Knowledge application deals with the process of using knowledge to make appropriate inferences. Conventional models of human problem solving assume that knowledge is organized into situation-action rules. Thus knowledge application is a matter of matching situations to actions. A rule-based model such as MYCIN's (a medical diagnostic program) production system (Shortliffe 1976) is a typical example. However, a rule-based model is rather weak in explaining the inferential processes involved in diagnosing. Davis (1983) observed that sometimes the diagnostic reasoning may come "from first principles", i.e., from an understanding of causality of the device being examined. Several

researchers incorporate this approach into the knowledge structure such that more sophisticated application of knowledge can be explained. Chandrasekaran and Mittal (1983) proposed a diagnostic problem-solving structure which has all the aspects of the underlying knowledge needed for diagnostic reasoning "compiled" into it. de Kleer and Brown (1983) studied mental models of devices and tried to derive inferences from those models. These approaches try to organize the relevant system knowledge in specific ways so as to aid the inferential process. Research along these lines is expected to benefit the development of intelligent computer aids for training. Chapter VI describes a modeling effort towards this goal.

CHAPTER VI

A MODEL OF FAULT DIAGNOSIS PERFORMANCE

From the experimental results discussed in the preceeding chapters it was seen that there was significant difference in the sequences of actions taken by subjects. This could possibly be the result of individual differences. Everything inexplicable by the proposed hypotheses or the central tendency could perhaps be classified and then discarded as individual differences. This may be adequate if the central tendency is overwhelming. However, since there are no obvious patterns and a strong central tendency is absent, a different explanation is needed. The modelling approach discussed in this chapter provides such an explanation.

Models of fault diagnosis performance can be classified into two types: macro and micro. Macro models describe performance strategies or behavior patterns. They describe general problem-solving rules that are abstracted from observations of subjects' diagnostic behavior. A model proposed by Rouse and Hunt (1984) that uses topographical rules (T-rules) and symptomatic rules (S-rules) is a typical example. These rules were employed successfully to classify diagnostic behavior. They explained "what" has been performed, but not "how" and "why" a particular action was chosen or formed. Rouse and Hunt also proposed a fuzzy set approach to explain how a particular rule is chosen in terms of the intersection of fuzzy sets of recalled, applicable, useful, and simple rules. However, no effort was made to explore how system knowledge was

acquired or why different weights were associated with different alternatives.

Contrasted with macro models, micro models consider individual differences. Therefore, a micro model should specify the formation of knowledge and the mechanisms behind actions. Most AI models can be classified as micro models since they usually deal with the knowledge representation and retrieval problems in a precise way. For example, Schank's (1982) approach to Memory Organization Package (MOP) not only specifies how memory is organized, but also describes the mechanism of reminding. Kolodner (1983a, 1983b) proposed a computer model of long term memory. Later she applied this model to medical diagnosis (Kolodner and Kolodner 1984) which demonstrated experience's role in acquiring expertise. She showed how medical knowledge is learned, modified and generalized. Medical diagnosis differs from engineering diagnosis in one important characteristic: medical knowledge about how the body functions is less structured than engineering knowledge about how a system works. Therefore, there is no clearly defined model to explain the patient's body function. An engineering diagnosis task is definitely better defined and structured than a psychiatric diagnosis.

In this chapter, a micro model is proposed that deals with knowledge formation and the mechanism of fault diagnosis behavior in a marine powerplant. Such a model could help identify essential components of fault diagnosis behavior and hence choose appropriate levels of simulator fidelity for training. A micro model might also reveal the limitations of human operators, and help develop better aiding schemes.

Kuipers and Kassirer (1984) pointed out that in building an intelligent computer system which deals with causal reasoning, two types of constraint have to be considered: computational constraints on knowledge representation, and empirical constraints on expert's reasoning process. Computational constraints require that a knowledge representation be computer implementable. Empirical constraints may direct the formation of the knowledge representation in a format that reflects how experts solve problems.

The model discussed below is based on observations of real fault diagnosis behavior. It tries to incorporate all data and protocols into a coherent model. The model considers empirical constraints only. We discuss two types of knowledge and their structures. Then the characteristics of diagnostic behavior on the marine powerplant are summarized. A conceptual entity called an "hypothesis frame" is employed to account for the observed behaviors. Finally, an integrated diagnosis model taking into consideration these two types of knowledge structures and the hypothesis frame is described.

This model is not intended to be a robust theory of diagnosis. As noted by Williams et al., (1983), in their paper describing mental models of a simple physical system,

"We do not claim that the models represent a theory of reasoning nor the simulations a test of a theory. Rather, the models are employed as tools to help us structure our description of the subjects' behavior." (pp. 136)

This chapter tries to structure the observations from subjects' behavior

in real diagnosis tasks.

Knowledge Representation

Based on observations and protocols (see Appendix C) taken from the experiments, two types of knowledge were identified: symptom knowledge and system knowledge.

Symptom Knowledge

Troubleshooters accumulate handy rules through years of experience. These rules associate obvious symptoms with non-obvious symptoms or causes. In the experiments some actions were generated fairly quickly, suggesting a direct association between symptoms and causes. For example, given "fluctuation of fuel oil pressure," most engineers immediately diagnosed that there was water in the fuel.

Theoretically, all symptoms can be derived mentally, given an appropriate representation of the system dynamics. However, some symptom derivation processes may be slow and error-prone due to human cognitive limitations. Symptoms that indicate environmental changes are good examples. Knowing the dynamics of the system might not help distinguish the causes for noises of different intensity and frequency. Experience plays an important role in associating causes with environmental symptoms such as noise, vibration, smoke, pressure fluctuation, etc.

The direct association between symptoms and causes can be represented with production rules in the IF-THEN format. When diagnos-

ing a system failure, subjects match the known symptoms with the antecedents of the rules. If matches are found, the action parts of the rules will be triggered. Causes associated with a particular symptom seem to have high generality across subjects. For example, rule 1 in Table 6-1 lists the most popular causes that are associated with the symptom "vacuum low at reduced speed". Gland seal steam was referred to by 20 out of 27 subjects in experiment two through six. Cooling water system and condensate pump system were referred to by 23 and 18 out of 28 subjects respectively.

There are three characteristics of these rules that make this matching process efficient.

1. Rules may be partitioned into groups such as steam-related group, noise-related group or fuel-related group, etc. This grouping scheme makes the rule-matching process efficient. A few examples of these symptom-cause rules are listed in Table 6-1.
2. The relationship between symptoms and causes is an inverted tree such as represented in Figure 5-5. Symptom-cause branches are prioritized. The priorities may change as experience accumulates and differs a lot from subject to subject. The priorities are used to guide the generation and evaluation of hypotheses. An example of a priority list can be found in the rules of Table 6-1 which is derived from the data collected.

Table 6-1: Examples of symptom-cause rules. The actions for a given condition are shown in decreasing order of priority.

Rule 1:

IF vacuum low and reducing speed

- THEN
1. check gland seal regulator failure frame
 2. check circulating water failure frame
 3. check the condensate pump failure frame

Rule 2:

IF vacuum low at full speed

- THEN
1. check condensate pump failure frame
 2. check vacuum pump failure frame
 3. check condenser failure frame
 4. check circulating water failure frame
 5. check gland seal regulator failure frame
-

3. Rules with a large number of conditions to satisfy are tighter than those with a small number. A tighter rule is usually preferred because it means a smaller and more accurate searching space for possible causes. For example, rule 1 in Table 6-1 is tighter than rule 2 in the same table. Therefore, rule 1 is preferred if both conditions are satisfied.

Another type of symptom knowledge is "facts" about the specific system being diagnosed. Facts are information about a particular system. An example of a fact is "This ship has been notorious on her atom-

izing steam system". Facts describe general observations of a particular system without clear association with any causes. Facts can help reduce the effort on the generation and evaluation of hypotheses. However, these are very system-specific. The symptom-cause pairings remain the same when the troubleshooter is switched to other ships, while the facts may be unique for a particular system.

System Knowledge

An engineering system is relatively well defined. Although there was no clear evidence of a homogeneous mental model about the system across subjects, there appear to exist hierarchical models of subsystems and components. Knowledge about the system is organized at several abstraction levels. In general, modules at higher abstraction levels are more function-oriented, while modules at lower abstraction levels have clear physical correspondence (Rasmussen 1985). However, at lower levels, the relationships are not entirely hierarchical since certain components and subsystems may be common to more than one system (Govindaraj 1984). For example, a steam system module may be defined as heat transfer unit, while a boiler itself may be defined as a collection of its major physical components. The following describes part of a possible hierarchical model of the powerplant.

The highest abstraction level may contain only the major subsystems such as steam, condenser, fuel, lube oil, auxiliary steam and turbine systems. A lower abstraction level in the steam system may contain the furnace, boiler, economizer, and air supply system. A possible lower abstraction level for the boiler may contain the superheater, the

desuperheater, the attemperator, the boiler drum and the boiler tubes. An even lower abstraction level for the superheater may contain the first pass pipe, the second pass pipe, the third pass pipe, the fourth pass pipe and the fifth pass pipe. A hypothetical hierarchical tree is shown in Figure 6-1.

Each module in this hierarchical tree may contain three types of information: definition, function, and connections. Table 6-2 shows a hypothetical boiler system module. In this boiler system module, there is no need to mention the details of the furnace system such as how the fuel gets mixed with the air or the role of air registers, etc. If in the process of diagnosing, more details about any submodules in the lower abstraction level is needed, the submodule can simply be accessed through appropriate links.

The definition and function of a module is described in terms of its descendants in the immediate lower abstraction level or modules on other branches of the tree. Links to modules of other branches of the tree denote the connections among modules with different functions. For example, the condenser may be functionally classified as a part of the cooling system. However, there is a connection between the auxiliary steam system and the condenser through the 35 lb steam line. This connection is shown on Figure 6-1 by the dashed line connecting the condenser and the auxiliary steam system.

When subjects were asked to describe how the system works in terms of higher abstraction levels, almost everyone described it in a similar fashion. This indicated that subjects at least had a consistent

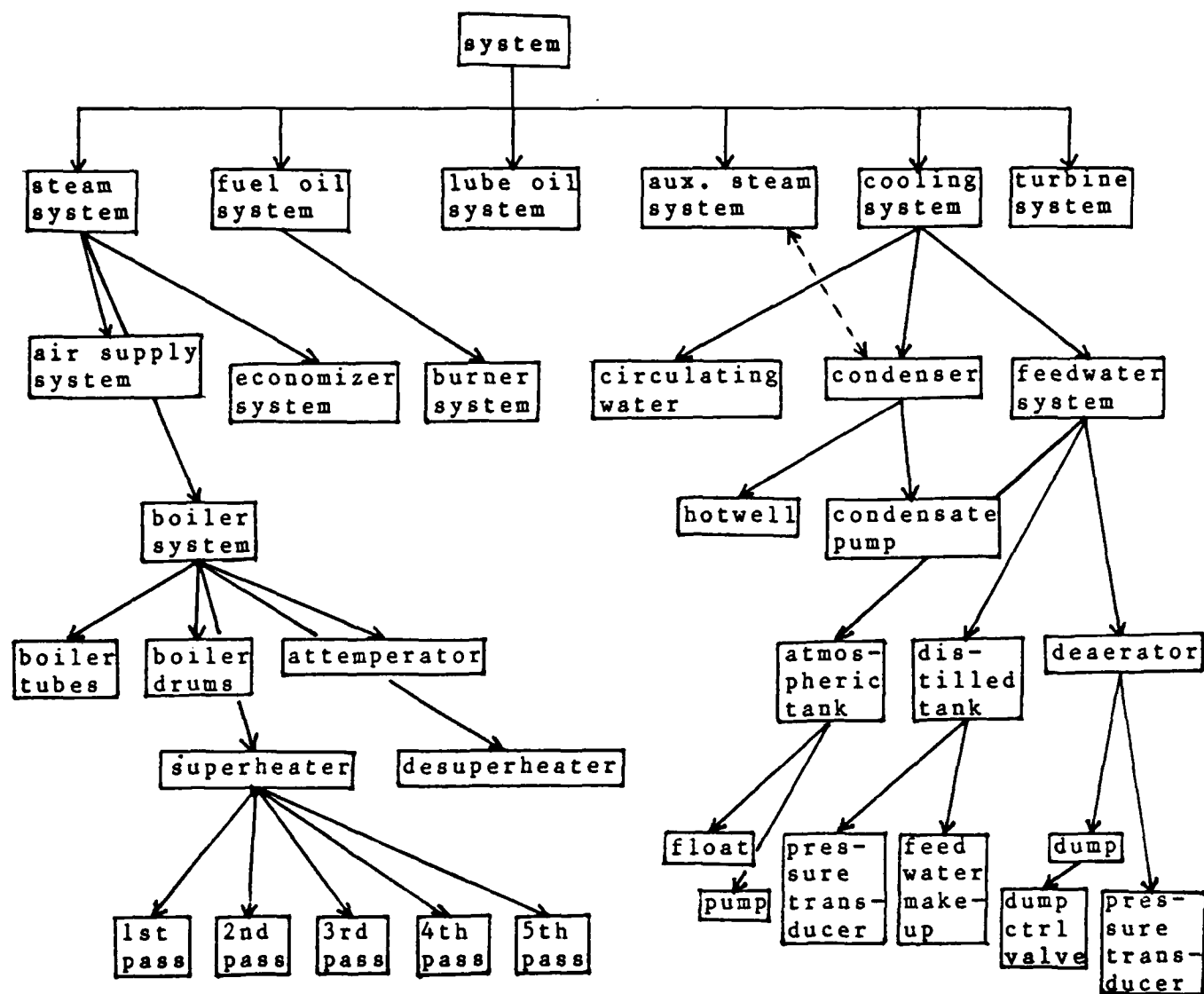


Figure 6-1: A hypothetical hierarchical tree

Table 6-2: The Boiler System Module

Definition:

A boiler system transforms the heated water into steam.

Function:

Feedwater goes through economizer to be heated before being fed into the boiler.
The boiler is heated by the furnace which mixes air and fuel oil for burning.

Connections:

Inputs are feedwater from feed system, fuel oil from fuel oil supply, and air from air supply.
Outputs are steam to steam system, gas and smoke to stacks.

view about the powerplant at the higher abstraction levels. However, the diagnostic process varied from subject to subject. Here are some observations about the diagnostic process.

Apparent Characteristics of the Diagnostic Process

1. Diagnosis proceeds in a hierarchical manner. Given symptoms, subjects start reasoning at a higher abstraction level and generate hypotheses using available information. Hypotheses generated in a higher level set the direction of diagnosis for a lower abstraction level. The following excerpt from a protocol showed this hierarchi-

AD-A161 361

MODELING FAULT DIAGNOSIS PERFORMANCE ON A MARINE
POWERPLANT SIMULATOR(U) GEORGIA INST OF TECH ATLANTA
SCHOOL OF INDUSTRIAL AND SYSTEMS. Y D SU AUG 85
GIT-TR-85-3 N00014-82-K-0487

2/2

UNCLASSIFIED

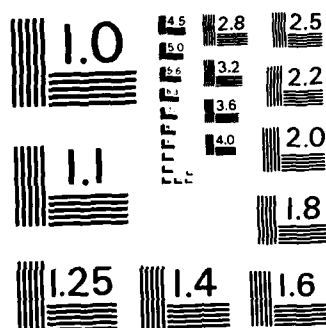
F/G 13/10

NL

END

FILED

SEP



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

cal approach.

L217: Need to find out whether there is too much demand for the boiler,

L218: Whether the boiler pressure is dropping in the auxiliary system or the main system.

L219: In order to bring the pressure back, you have to reduce the load.

L220: We have to come back to a reduced speed,

L221: To secure some other steam demand and find out the problem.

(Experimenter intervention to keep conversation going)

L223: You have to check to see if both burners in both boilers are operating properly.

L224: If the fuel oil is following the demand and the air,

L225: If there is not enough air we won't get proper atomization. I mean proper burning.

L226: Have to check different pressure.

L227: Whether the temperature of the boiler is OK

L228: To see if it would have dirty tubes.

The subject first diagnosed the problem in terms of functional concepts like "too much demand on the boiler" (L217). Following this diagnosis of a high abstraction level, the subject suggested some remedial actions (L219, L220, L221). Then after assuring that the boiler was not overloaded, the subject proceeded to diagnose the problem further using "too much demand on the boiler" as a guide-

line. He shifted attention to a gross but concrete component "burner" (L223). To identify what was wrong with the burners, he moved down the hierarchy tree to the fuel line and air line inside a burner (L224, L225). At this moment, he checked the system state to find out the status of the fuel and air lines going into the burners. They were alright; the subject therefore came back to a higher abstraction level to resume another line of reasoning (L226, L227).

2. Diagnosis proceeds with backward and forward reasoning. It is commonly known that, given symptoms, subjects search for symptom-cause pairs to proceed with the diagnostic process. This is a backward reasoning process which goes from symptoms to causes. However, it is observed that subjects also use forward reasoning to help formulate and eliminate hypotheses. Forward reasoning process allows the subjects to reason through the system dynamics over time starting from a component of interest. For example, a subject was observed to reason in the following manner.

L530: If cold water goes into the feedpump, it would cause the feedpump to flash the steam, and the pump would then lose suction.

In this case the subject mentally derived how the system would react if cold water goes into the feedpump. This is a forward reasoning process.

3. The symbolic descriptions of quantities are stated in qualitative terms. For example, high pressure, low vacuum, enough flow etc., are used to describe the system variables. Therefore, qualitative state descriptions seem adequate.
4. Knowledge appears to be chunked: relevant symptoms, inferences and structures are grouped together mentally for easy access. Subjects showed great efficiency in evolving through system states qualitatively. For example, when asked about what would happen if the 35 lb line had problems, the subject responded quickly with the dialogue from line L135 through line L138 in Table 6-3.

The symptoms (hearing the 35 lb make up, hearing the feedpump speed up a little), the inferences (from 35 lb failure to the make-up, to the feedpump state change, to the dumping into condenser), and the structures/connections (the relationships between 35 lb line, feedpump, high pressure steam, low pressure steam, and the condenser), are clearly stated.

Hypothesis Frame

These four observations lead to the construction of a useful conceptual entity called hypothesis frame. A frame as described by Minsky (1975) is a general way of representing common knowledge. Frames contain information about many aspects of the objects or situations they describe. The "hypothesis frame" or prototype has been used for organizing disease types in medical diagnosis research. CENTAUR (Aikins, 1979) and PIP (Szolovits and Pauker 1978) used "hypothesis frame" as the

Table 6-3: Protocols from subject S31 while solving Failure 19. Lines beginning with "E" were interruptions by the experimenter to keep the onversation going. Lines beginning with "L" were uttered by S31.

Failure 19 (Vacuum drops at reduced speed) (S31)

L100: First indication is a lot of vacuum loss.
 L101: So, first thing I'll check will be the gland seal
 L102: (to see) if it has pressure after regulator.
 L103: You know on the board it is at least, say, 1 or 2 pounds.
 L104: If that is OK, I believe, you check circulator maybe,
 L105: to see that you have enough water going back into the condenser
 L106: to keep the air ejector cool.
 E107: Why the air ejector?
 L108: Oh! If you talk about the vacuum pump... I am not quite familiar with the vacuum pump. Let us say the
 L109: I know with the air ejector you lose vacuum that way.
 L110: What else can be wrong?
 E111: So, those are most common problems?
 L112: Yes, I would think so, unless you...
 L113: Assuming the condensate pump running normally,
 L114: that will be the first indication, the major cause.
 E115: So, what else?
 L116: Check level of condenser.
 E117: What do you want to find?
 L118: You want to find out if it is normal,
 L119: if it is too high
 then... you see it could be getting water from somewhere else.
 E120: What do you mean by somewhere else?
 L121: Say, for example, condensate pump may not be functioning properly.
 E122: Think aloud.
 L123: It would be ..., well.
 If the level is OK, then you look for something else.
 L124: We just slow down.
 L125: You assume everything else is OK.
 L126: In other words, circulating water is fine,
 L127: no reason for that, which you think probably will be ...
 L128: I don't think it has any problem with the bleeds
 L129: I don't know. The symptom is...
 E130: Vacuum dropping when slowing down to 50 from full speed.
 L131: Possibly the bleeds, you could lose it that way.
 E132: Any other possibilities?
 L133: Like I said, the dump could cause it,
 L134: when you have problems with the 35 lb line.
 L135: If that was the problem you hear the 35 lb make up
 and you hear the feedpump speed up a little.
 L136: You definitely hear the 35 lb line making up
 L137: from the high pressure steam coming into the low pressure steam
 L138: where the low pressure dumping it into the condenser.
 (S31 checked the 35 lb dump regulator and found the solution)

main construct for representing disease categories.

An application of the frame concept to fault diagnosis involves assuming the existence of a "hypothesis frame" which contains all relevant information about a particular failure situation. In other words, for each possible failure or hypothesis there is a frame associated with it. Each hypothesis frame may contain four slots: symptom slot, component slot, inference slot, and flow slot. Table 6-4 shows an example of hypothesis frame for condensate pump failure. Six different symptoms are associated with this failure. These point to proper rules in the symptom knowledge base. Eight relevant components are listed. These point to modules in the system knowledge base where the relationships among the components are placed in proper abstraction level. Components may point to modules of higher abstraction levels. For example, the atmospheric feed tank may point to the cooling system module (see Figure 6-1) instead of atmospheric tank module or feedwater system module. This scheme facilitates the diagnostic process that involves reasoning from higher abstraction levels and proceeding in a hierarchical manner. The inference slot contains relevant inferences which may be derived from the system knowledge base. The flow slot indicates which type of flow is involved.

The relationships among system knowledge, symptom knowledge, hypothesis frame database, and known symptoms are depicted in Figure 6-2. Known symptom set is used to match rules in the symptom knowledge base. Rules matched cause the appropriate frames from the hypothesis frame database to be chosen. New symptoms may be found while processing

Table 6-4: Hypothesis frame for a condensate pump failure

Symptom slot: (All are indexed back to proper rules
in the symptom knowledge base.)

condensate pressure low
deaerator pressure low
deaerator level low
vacuum pressure low
condenser level high
LP turbine exhaust temp high

Component slot: (All are indexed into proper modules
in the system knowledge base.)

condensate recirculate control valve
feedwater make-up control valve
feedwater make-up regulator pressure transducer
atmospheric drain tank
deaerator feed tank
main condenser
condensate pump suction valve
condensate pump discharge valve

Inference slot:

1. Condensate pump failure will cause low condensate discharge pressure;
2. therefore more condensate remains in the condenser than normal.
3. This in turn causes condenser level to go high.
4. Once the level goes high, it loses vacuum and the LP turbine exhaust temperature increases.
5. low condensate discharge also causes the deaerator level and pressure to go low.
6. This in turn triggers the feedwater makeup control.
7. which results in the atmospheric drain tank dumping water into the deaerator tank.

Flow slot:

feedwater

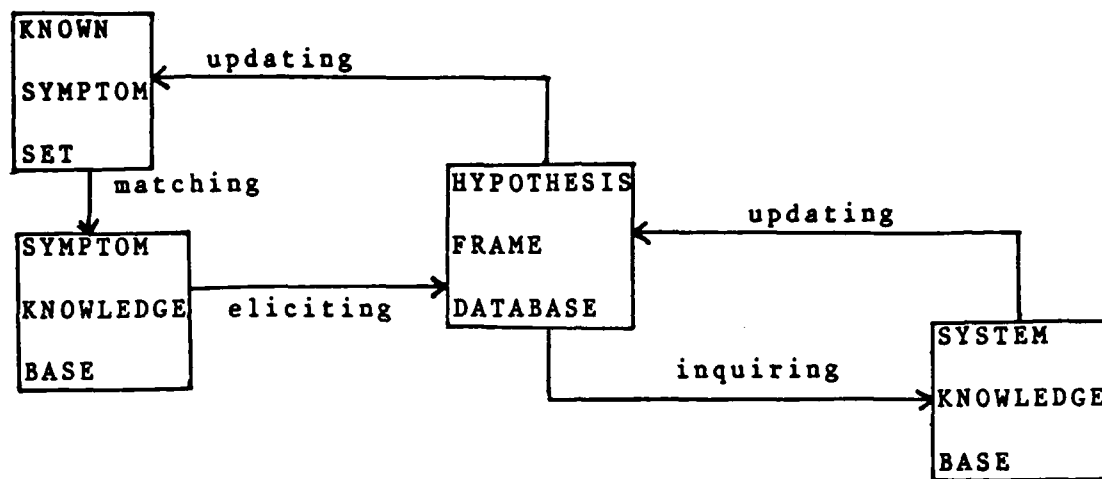


Figure 6-2: Relationships among known symptom set, symptom knowledge base, hypothesis frame, and system knowledge base. Known symptoms are used to for rules in the symptom knowledge base. Rules matched cause frames to be chosen. Known symptom set should be updated if new symptoms are found. The system knowledge base provides the information about system structure to support the processing of frames. Frames may be updated through inquiry with of system knowledge base.

frames. In this case the known symptom set should be updated. The system knowledge base provides the information about system structure to support the processing of frames. On the other hand, information in a frame may be updated (e.g., corrected or augmented) by interacting with the system knowledge base. Also new frames may be acquired through years of diagnostic experience. In other words, learning may be involved. However, the mechanism of acquisition and modification of

hypothesis frames is a complicated learning process. The learning process is not discussed in this thesis. Nevertheless the existence of hypothesis frames is assumed and a model for the diagnostic process is described based on this assumption.

A Model for the Diagnostic Process

At the beginning of a diagnostic process, there are some given symptoms in the known symptom set. As the diagnosis proceeds, this symptom set is updated. The updating process may not simply be the addition of new symptoms. Deductions may be drawn from symptoms. For example, while solving Failure 23 in which both feedpumps were heard to make noises, most subjects would conclude that the trouble lay upstream of the feedpumps. The trouble could not be the feedpumps themselves because it is very unlikely that both pumps would go down at the same time. This new deduction about the system from the symptoms becomes a part of the known symptom set.

Symptoms in this set are used to match with the symptom-cause rules. Rules in the symptom knowledge base may be organized into partitions of several different categories, such as, steam-related, noise-related, fuel-related, etc. There are two ways to make the matching process efficient. The first is the facts about the particular failure under investigation. These facts are generalized from the known symptom set and the experience of evaluating frames. Examples of commonly used facts are "trouble should be upstream of the feedpumps since both pumps

have failed", or "it is a steam demand problem", etc. These facts form a search direction that guides the rule-matching process to search in certain partitions only instead of the whole database.

Another way to speed up the matching process is to use heuristics. An example of a heuristic is: "choose frames of higher priority". Rules in the symptom knowledge base are prioritized. Subjects have a tendency to choose the higher priority rules. In Table 6-3, given vacuum loss when slowing down the ship (Failure 19), S31 checked the gland seal steam system first, followed by the 35 lb line system. In fact, 13 out of 19 subjects observed checked gland seal before they checked the 35 lb line system. Another example of a heuristic is: "choose tighter rules whenever possible". It was observed that if the vacuum loss symptom occurred at full speed (Failure 3), a different sequence of actions would be adopted. 19 out of 20 subjects checked the condensate system first instead of the gland seal system which was the most preferred action when vacuum loss was paired with speed reduction.

Once an appropriate rule is found relevant, the corresponding hypothesis frame is proposed. During the processing of a frame, i.e., evaluating the hypothesis, subjects interrogate the system to match the information in the slots. Any new symptoms found are fed back to the known symptom set. If new symptoms are found in the process of evaluating the hypothesis frame, these new symptoms and old symptoms may be used to search in the symptom knowledge base for tighter or higher prioritized hypothesis frames. However, if inferences drawn from frames are in conflict with observations from the known symptoms, then these

frames should be discarded.

If no new symptoms or rules are found, subjects may resort to another type of heuristic. Some observed heuristics are given below.

1. Physical closeness: If subjects failed to find any significant symptom from the system, they tended to search through subsystems or components that were physically close to the subsystem/component that they currently hypothesized to have failed. For example, while solving Failure 7 in which the HP turbine vibration excessive alarm sounded and the engine noise level was higher than normal, S21 started from the superheated steam frame. Then he checked every component in sequence, starting from the burner and going all the way up to the FD fan. All these components were physically close, but not necessarily relevant to this particular failure. Among 6 subjects who had difficulty with this failure, 4 had a tendency to process subsystems that were physically close.
2. Pick a best-matched frame: Investigate frames that are matched best. In Table 6-3, after failing to confirm the gland seal and circulating water frames, the subject tried to process the bleed frame. However, at first he was not convinced that bleed was the problem (L128). The reason was that there should have been other symptoms preceding the loss of vacuum (L135) if bleed was the problem. Nevertheless this frame was chosen under the best-matched heuristic.

3. Direct mapping from symptoms: Investigate frames that contain the components or subsystems indicated in the symptoms. For example, 8 out of 21 subjects picked feedpump failure frame when presented with Failure 8 in which feedpump noises are heard.

Table 6-5 lists types of heuristics used by subjects who had difficulty solving failures rated as moderate and difficult in Chapter V. Figure 6-3 presents this model in a flowchart-like schematic.

Table 6-5: Types of heuristics used by subjects. PC denotes Physical Closeness. BM denotes Best-Matched heuristics. DM denotes Direct Mapping from symptoms. PC and DM were identified from the subjects' recorded actions, while BM was identified from protocols which were only conducted on some subjects. Hence BM occurred less often than PC and DM. Note that a subject might use more than one type of heuristic in a failure. Also note that not every heuristic is identified in this analysis.

	Failure	PC	BM	DM	no. of subjects
moderate	16	1	-	7	8
	20	7	-	3	10
	21	5	-	1	6
	26	9	-	-	9
difficult	7	4	-	-	6
	8	3	1	8	21
	9	13	-	2	14
	10	11	-	2	15
	19	6	1	2	8

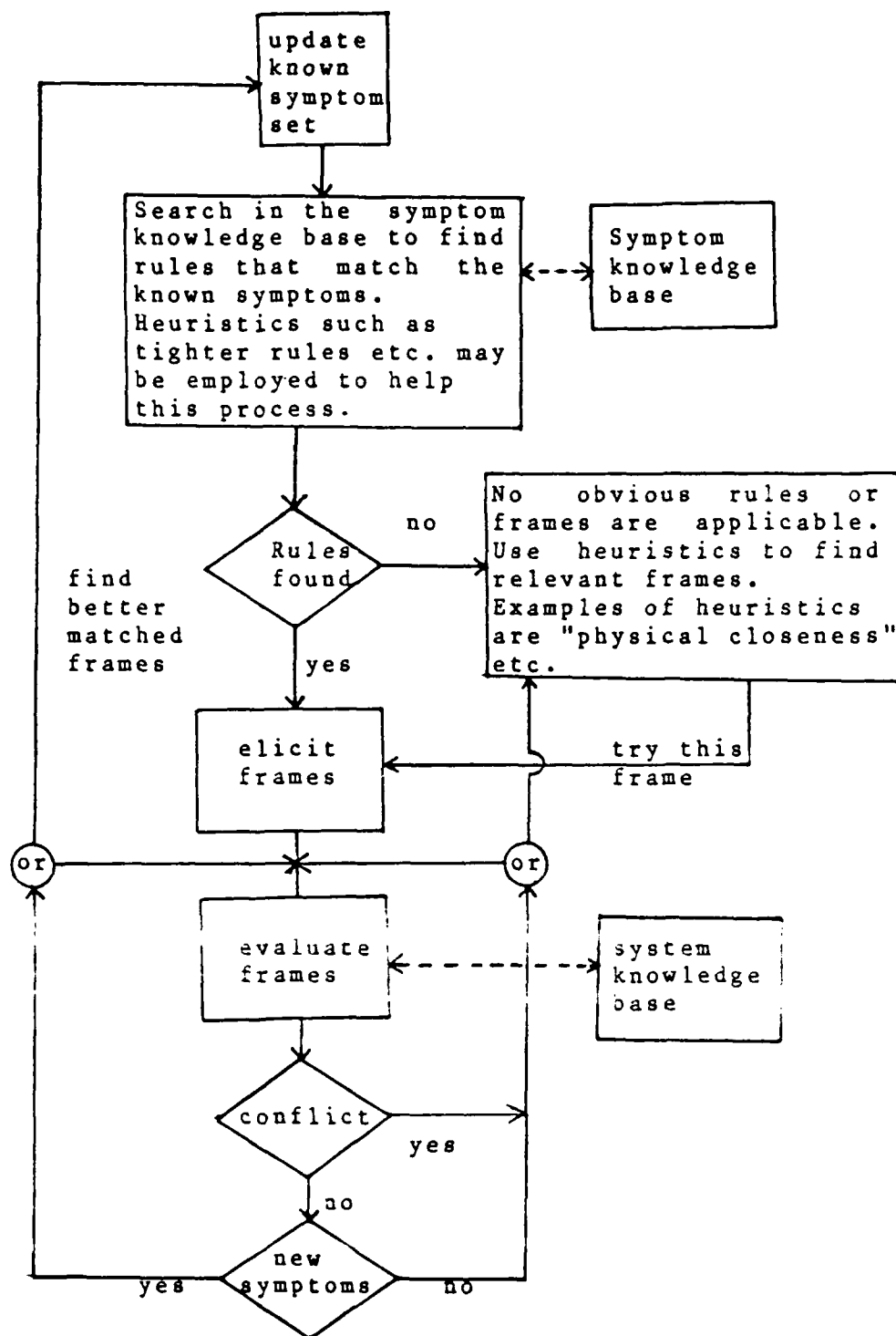


Figure 6-3: Flowchart of the diagnostic process

Interpretation of Protocols

The model proposed in previous sections is abstracted from experimental data and protocols. There are six pieces of protocols taken from four subjects (S31, S32, S33, S34). Four out of six protocols were taken while subjects were solving Failures 8, 9, 19 and 23 respectively. The other two recorded subjects' description of how the powerplant works, and the function of gland seal steam. Four diagnostic protocols are interpreted below in terms of the model outlined in the last section.

Protocol 1:

Table 6-3 lists an excerpt from a protocol taken while S31 was diagnosing Failure 19 in which vacuum was dropping at reduced speed. S31 might take the vacuum dropping as a symptom. However, "at reduced speed" seemed to provide another piece of information. Therefore a tighter rule that considered both was preferred. He matched antecedents of the rules in the symptom knowledge base with these two known symptoms. If he had the rules as in Table 6-1 in his repertoire, then rule 1 instead of rule 2 would be matched. Therefore, he first processed the gland seal regulator failure frame since this frame had the highest priority (L101, L102). Once the frame was brought into attention, he tried to match and evaluate the frame by interrogating the system about the information in the slots. This was a straightforward failure in which the gland seal steam pressure had to be low. So S31 checked this information with the system. However, the system responded that the gland seal pressure was normal. The subject rejected this frame without

finding any new symptoms.

He kept searching down the priority list and processed the second most probable frame, the circulating water failure frame (L104, L105). He checked the system to see if the circulating water pressure was low, which might be the first symptom in the symptom slot. The system responded that it was alright. Therefore he processed the third frame in the priority list --- the condensate pump failure frame (L112, L113). This time, instead of checking symptoms in the symptom slot, he followed the inference path in the inference slot (see Table 6-4) to reach an assertion that if the frame was appropriate then the condenser level must have been high (L118, L119). So he checked it and found that it was alright again. At this moment S31 was stuck and found no further symptoms to help continue the diagnostic process.

After a little pause, he seemed to pick up a frame that he thought was related but did not match the known symptoms well (L128). Here we could assume that he was applying the "pick a best-matched frame" heuristic. By using the information provided in the inference slot of the new frame (the 35 lb line system failure frame), he was able to make several qualitative inferences about the system changes quickly (L135, L136, L137). He checked the 35 lb steam dump regulator and found the solution.

Protocol 2:

Another protocol was taken while S33 was solving Failure 9. It showed how heuristics were employed and how frames with conflicting

symptoms were rejected as well as how modules in the system knowledge base were employed to explain failures. Table 6-6 presents an abridged version of the protocol.

Failure 9 described a situation in which white haze was seen. S33 quickly responded that too much air was the problem (L552). After locating the failure in a rather high abstraction level (in this case ---- the air supply system), S33 got into a lower abstraction level reasoning starting with the FD fan (L553). At this moment the FD fan failure frame was brought into attention. The inference slot of this frame might contain the statements such as L559 and L560. These inferences led S33 to conclude that air pressure should be changed. From L554 to L558, S33 tried to explain how the control of the FD fan works in terms of modules of lower abstraction level such as diaphragms, and control signals. In fact, a rather simplified model of how the fan functions along with the boiler was implied. Figure 6-4 depicts this model in a block diagram.

This type of knowledge does not appear in the hypothesis frame itself. However, the component slot of the frame might point back to modules in the system knowledge base in which this type of knowledge appears as definitions, functions or connections with other modules.

Having failed to find the fault in the air supply system, S33 decided to try the fuel oil supply system. He found that fuel oil temperature was normal and pressure was high (L575, L576). He decided to investigate the fuel oil system. However, in reality most failures in

Table 6-6: Protocols from S33 while solving Failure 9.

Failure 9 (White haze is seen on STBD boiler periscope) (S33)

L550: White haze means too much air.
 L551: Let us check the STBD burners' management system.
 It says it is running.
 L552: Look at the air system.
 Too much air may be the problem.
 L553: So burners' management sytem is OK.
 Let us start with the fan and work our way down.
 L554: I got pressure to the boiler out of the fan.
 L555: Signals go into the combustion control of that fan,
 L556: that regulates the fan.
 L557: Maybe the combustion control is sending a wrong signal
 to the top of the diaphragm.
 L558: Or the diaphragm is broken.
 L559: If it fails open then I have more air than I need.
 L560: So check the air pressure.
 L561: It is high.
 L562: The fan is OK since it is running.
 Maybe the crossover..... normal.
 L563: Go down the line see if I can pick up something else.
 L564: Check windbox down the line.
 It says high..... but crossover is normal.
 L565: So nothing is wrong along this line.
 FD fan is running.
 L566: Maybe it has to be slow. Well, OK.
 L567: I will say maybe my fuel oil has problem.

 L575: Let us check out the temp of the fuel.
 L576: Temp is normal, pressure is slightly high.

 L586: Check the strainer... no, that cannot be it.
 L587: Because that should cause pressure to go low.

 L592: Something wrong in the pump? No, it cannot be because
 it would be low too.
 L593: If the relief valve is open....
 L594: It would be low too.

 L605: I am stuck.
 (The solution should be blockage in the STBD boiler tubes)

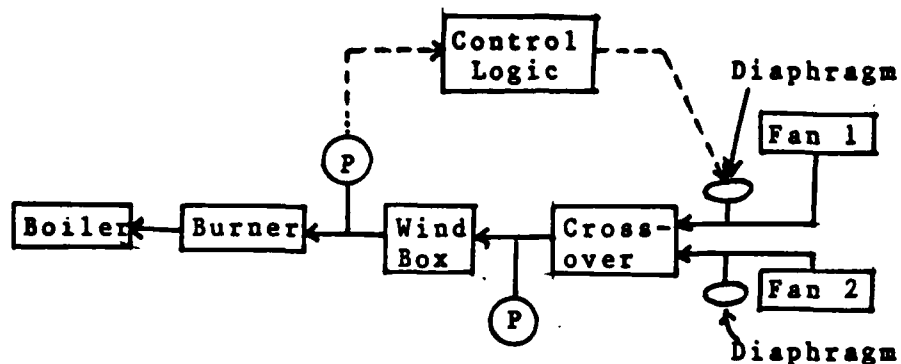


Figure 6-4: A simplified model of the air supply system which was implied by S33 in his protocol.

the fuel oil system would show opposite symptoms, i.e., low instead of high fuel oil pressure. "Direct mapping from symptoms" heuristic might be employed to explain why S33 took this move which turned out to be futile.

S33 then picked up component failure frames along the fuel oil line and found that they all conflicted with the fact that fuel oil pressure was high (L586, L587, L592, L593, L594). So these frames were rejected one by one. Finally, S33 claimed that he was stuck and gave up (L605). The key to the solution lies in the rules associated with white haze. S33 seemed to have only one cause to explain for white haze --- too much air. However, in this case, a boiler tube leak was the cause of the white haze. The high air pressure and fuel oil pressure were the consequences for compensating for steam loss.

Protocol 3:

Table 6-7 shows the first part of the protocols taken while S32 was solving Failure 23. In failure 23 boiler pressure dropped and smoke was seen on both boilers while the ship was brought back to full speed ahead. At the same time low atomizing steam alarm sounded. S32 first considered only part of the known symptoms, i.e., both pressures dropping. He retrieved the fuel oil system failure frame and found conflicts of symptoms (L201). He then rejected this frame and processed another frame --- smoke frame. This time he successfully explained both pressure and smoking symptoms (L203 - L207). At this moment, low atomizing steam symptom was brought into attention. S32 was able to make several quick inferences about the failure of the atomizing steam regulator. The inferences seemed to match with the known symptoms.

S32 then proceeded to find the solution --- atomizing steam reducing regulator failure. This analysis showed a pitfall troubleshooters committed often: neglecting known symptoms. If S32 did not overlook the low atomizing steam from the beginning he need not have gone through the first two frames. This seemed to point out a possibility for aiding. An appropriately designed fault diagnosis aiding system should be able to help the troubleshooters choose the tightest rules by not overlooking symptoms. Rouse (1981) noted similar phenomenon. He found that valuable negative information was overlooked in a context-free logical troubleshooting task.

Table 6-7: Protocols taken from S32 while solving Failure 23.

Failure 23 (When rising back to full speed, boiler pressure dropped and smoking in both boilers. Also the low atomizing steam alarm sounded) (S32)

- L200: First we need to find out what's causing the boiler pressure to drop.
L201: Check fuel oil pressure.
You said it is OK.
L202: And you said smoking.
L203: That'll probably indicate the pressure would be extremely high
L204: because it is a large load from the boiler.
L205: So that could cause the smoke.
L206: It would be too much demand for the boiler to make up.
L207: That's the reason the boiler pressure is dropping.
L208: With low atomizing steam we have to check the 150 lb system
to see if that is normal.
L209: If that's normal, we have to go to the atomizing steam regulator
to see that is functioning properly.
L210: If that fails, you have poor atomization which undoubtedly
will trip the boiler.
E211: Cause the smoke?
L212: Surely, because you have low atomizing steam,
L213: and you would have to slow down to bring the pressure back
so you don't lose the plant.
L214: Doing that will bring up the fuel oil pressure back.
L215: Remedy the atomizing steam somewhat or bypass it and find out the fault.
That seems to be the problem.
(S32 found the solution: atomizing steam reducing regulator failure)
-

Table 6-8: Protocols taken from S34 while solving Failure 8

Failure 8 (Vibration, noise and speed fluctuation occurred in the no. 1 pump. When the engineer started the no. 2 pump, the same symptoms occurred again.) (S34)

- L450: Well, maybe loss of suction because that is common to both pumps.
 L451: Check the level of the deaerator (normal).
 L452: I have checked out the suction side of the water side of both pumps.
 L456: What will cause fluctuation. Maybe there is
 L457: I will go to my steam system.
 L458: Now I have eliminated the suction water side of the pump.
 I am up to the economizer.
 L459: Check temp gauges coming in and out of economizer.
 L460: I wonder if there is a leakage in the economizer?
 L461: Maybe the economizer is leaking.
 L462: Sometimes when you started up a pump it will cause speed fluctuation and trip out.
 L463: Inlet is normal; check the outlet.
 L464: If outlet is low, then that means the economizer is plugged up.
 L465: Normal. So every thing on the water side is OK.
 L466: Now I am going to the steam side of the feedpump.
 L468: Check if 35 lb line has back pressure....
 L469: Check if 35 lb pressure is normal (normal).
 L470: If that is normal, back pressure cannot be low.
 L471: I will be getting.... Pumps would not be fluctuating.
 L472: Now go on to the steam side of the inlet.
 L473: I cannot see nothing that is common...
 L475: Desuperheated steam for my feedpump....
 L476: Check if both of them come out right.
 L477: I want to find a cause common to both feedpumps.
 L478: I have checked my pressure and levels.
 L479: I wonder about if the temp of the steam is relevant...
 L480: If cold water goes into the feedpump,
 L481: It would cause it to flash the steam
 L482: And the pumps would lose suction.
 L483: A lot of times when we have problems with our deaerator
 L484: And we have to bypass the deaerator.
 L485: We take water from our distilled tank.
 L486: The water in the deaerator is 275 degrees,
 L487: And the water coming from the distilled tank is 90 degrees.
 L488: Difference in change will cause it to flash.
 L489: And the pump will lose suction.
 L490: I am going to check out the temp of the water coming from the deaerator.
 L499: Steam temp may be a problem.
 L500: 20 degrees low coming out of the desuperheater.
 L501: Temp of steam could cause fluctuation of the feedpump
 L502: Because the higher the temp the more the work you get out of the steam.
 L503: So if it is 20 degrees low, it got to be something wrong in my desuperheater.
 L504: I am now backtracking.... coming out of the superheater

L505: Going into the desuperheater.
L506: Backtrack a little bit further.
L507: Come out of the drum going to my superheater.
L508: I am going to check my drum level (normal).
L509: So, back into my superheater.
L510: It has to be the attemperator, the desuperheater or the superheater.
(At this point S33 found the solution ---- desuperheater failure)

Protocol 4:

Table 6-8 is an abridged version of the protocols taken while S34 was solving Failure 8. In Failure 8, vibration, noise and speed fluctuation occurred in the no. 1 feedpump. When the engineer started the no. 2 pump, the same symptoms occurred immediately. S34 seemed to be directed by the heuristics of "direct mapping from symptoms". The frame of feedpump failure was investigated first (L452). The known symptoms seemed to be augmented by an observation deduced from given symptoms, i.e., the fault must be common to both pumps (L450). Having failed to find any significant evidence to support this frame, S34 proceeded to the economizer failure frame (L458 - L466). This transition of frame involves no new symptoms. S34 did not seem to have any rule that could cover the symptoms (L456). Transition from feedpump failure frame to economizer failure frame may be explained by heuristics such as "physical closeness" or "best-matched frame". The economizer frame seemed to support the fluctuation symptoms. Several inferences were drawn from this frame to come up with a scheme for evaluation, i.e., test to see if differences existed between inlet and outlet pressures. Having failed to find evidence to support the economizer frame (L465), S34 considered the 35 lb line system failure frame (L466 - L471). However, the inference S34 drawn from this frame was in conflict with the known symptoms

--- fluctuation (L471). So, this frame was rejected. Then S34 went back to the feedpump failure frame again (L475 - L482). This time he investigated the possibility of cold water going into the feedpump. He found that it could cause loss of suction. This inference resulted in consideration of the deaerator failure frame for processing only to find that it was normal (L483 - L490).

Although the last hypothesis frame failed, the possibility of difference of temperature seemed to lead S34 to have the idea of steam temperature difference (L499). A new symptom was then found, i.e., temperature was 20 degrees lower than normal coming out of the desuperheater (L500). The desuperheater failure frame was then brought in for processing (L501 - L510). Finally, S34 found the problem in the desuperheater.

These four protocols show how diagnostic process proceeds from symptoms to the faulty components. The model seemed to interpret the protocols reasonably well. The following discusses how the attributes of the model fit the protocols, and compares the model with other approaches.

Discussions

This model is characterized by three key aspects: heuristics, knowledge bases, and diagnostic processes. Heuristics are employed when no obvious frames can be found. Five heuristics were identified for this model. They are: physical closeness, best-matched, direct-mapping, tighter rule, and higher priority heuristics.

Knowledge bases contain all relevant domain knowledge. They are: the system knowledge, symptom knowledge, and hypothesis frame bases. System knowledge base contains modules of different abstraction levels, and is organized hierarchically. The definition, function, and connections of components may be found in the modules. The dynamics of a component is described in qualitative terms and represented in a module. This has the potential to allow the model to "reason from the first principles" which has been advocated by Davis (1983). The symptom knowledge base is in some sense an index file for the hypothesis frames. The keys are the matching symptoms. The symptom-cause pairs associate matching symptoms with appropriate hypothesis frames.

The hypothesis frame base contains frames for individual failures (see Table 6-4). This is similar to the "compiled" deep knowledge concept of Chandrasekaran and Mittal (1983). In their proposal, the knowledge needed for diagnostic reasoning is "compiled" into the system knowledge structure. This allows reasoning to proceed efficiently without the need to access deep knowledge such as those of the "first principles". However, "compiled" deep knowledge scheme limits the access of knowledge to certain paths that are reachable under the knowledge organization. The information in hypothesis frames can be used independent of reasoning paths since they are independent of the knowledge structure. Therefore they can be elicited from any point of the reasoning process. This database is highly modularized. Modification and addition of frames can be accomplished with minimal effort.

The inferences provided in the inference slot in a hypothesis

frame make the forward reasoning possible. Using information contained in the inference slot, the model can infer about the system state as a result of the fault such as "boiler pressure drop" etc. This enables the model to generate schemes for testing the frames of interest against the system state.

There are four types of processes involved in the model. They are: symptom updating, frame elicitation, frame evaluation, and hierarchical reasoning. These processes are widely recognized and implemented. Genesereth (1982) advocated a computer hardware fault diagnosis system using hierarchical design models. The program, called DART, first diagnoses the system at a high level of abstraction to determine the major subsystem in which the fault may lie. It then focuses on the next lower level and repeats until the fault is found.

Frame elicitation and frame evaluation are conceptually similar to hypothesis formation and hypothesis evaluation. Hypothesis formation and evaluation has been widely used in medical diagnosis system (Shortliffe 1976, Pople 1977).

The model proposed in this thesis integrates these attributes of diagnosis into a coherent structure. Protocol analysis in the previous section demonstrated how this model can be used to explain experts' fault diagnosis behavior. Table 6-9 shows the number of times each attribute of the model occurred in the protocols. Judging from the frequency of occurrence and coverage of the attributes, the model seemed to fit the protocols reasonably well.

This model incorporates the knowledge structure with processing schemes. Hence it has the potential to be expanded to implement an intelligent on-line coach for troubleshooting.

Another possible use of the model is to develop aiding systems for fault diagnosis. Slovic et al. (1977) pointed out that the effectiveness of aiding depends upon the ability of the user to decompose the overall decision problem into its constituent elements. This model shows the mechanisms underlying the diagnostic actions. The diagnostic behavior, if viewed as a decision task, can be decomposed into subprocesses. Hence, better aiding schemes could be developed by considering subprocesses and relationships between subprocesses.

A robust theory of fault diagnosis should consider the influence of detection and compensation in addition to diagnostic performance. The model proposed above is not a robust theory of fault diagnosis in this sense. However, it serves to organize the observations from subjects' behavior in real diagnosis tasks. Further research is needed to make this model applicable.

Table 6-9: Occurrence of the attributes in the protocols.
Each entry indicates the number of times a particular attribute of the model is observed in a protocol.

Attributes		Protocols			
		1	2	3	4
Heuristics	Physical closeness				1
	Best-matched	1			1
	Direct-mapping		1		
	Tighter rule	1			
	Higher priority	1	1	1	1
Knowledge Base	Symptom	1	1	1	1
	System		1	1	
	Hypothesis Frame Symptom	2		1	
	Component				1
	Inference	2		1	3
	Flow			1	
Processes	Symptom updating		1		2
	Frame elicitation	4	4	2	7
	Frame evaluation	4	4	2	7
	Hierarchical reasoning		2		

Conclusions

A model for the fault diagnosis process has been presented based on the data from experiments and protocols. A conceptual entity called hypothesis frame was employed to account for the observations from data and protocols. Instead of reasoning from the dynamics and structure of the system, relevant information about a failure is "compiled" in the frames. Therefore subjects could make quick and reliable inferences which would otherwise be impossible. Given symptoms of a failure, subjects would first try to match a frame from their repertoire of rules. Heuristics might be employed to make this matching process efficient. Once a frame was identified, subjects could use the information contained in the slots to make inferences or conduct evaluations. If more symptoms were found, subjects might use the new evidence to search for better hypotheses. If inferences drawn from frames are in conflict with known symptoms, these frames are discarded. When no obvious frames could be processed, subjects could search in the system or symptom knowledge base for relevant information under the guidance of heuristics.

This model was applied to Failure 19 (S31), Failure 9 (S33) Failure 8 (S32) and Failure 23 (S34) with good results. This model may be regarded as reasonable for organizing the experimental observations and protocols from fault diagnosis performance. Further refinements and testing are needed before it can be applied in training and/or the design of on-line coach system for fault diagnosis.

CHAPTER VII

CONCLUSIONS

In this thesis, we have described an investigation of fault diagnosis behavior of experts in a realistic environment where marine power-plant simulators were used. The initial feasible set of possible faulty components and the strategy subjects used to alternate between hypothesis formation and evaluation stages were found to affect the efficiency of performance. This finding seems to imply that both context-specific (initial feasible set) and context-free factors (transition strategy) are important for fault diagnosis training.

A micro model of fault diagnosis that incorporates system knowledge, symptom knowledge and hypothesis frames into a coherent structure was proposed and used to interpret protocols. The model fit the protocols reasonably well. This model may be expanded to implement an intelligent on-line coach for troubleshooting.

The research discussed in this thesis used a low fidelity simulator. A possible extension of this research is to study fault diagnosis behavior using moderate and high fidelity simulators as described in Chapter III. Both the moderate and high fidelity simulators have the system dynamics incorporated, which enables the presentation of a process view of the system instead of a frozen picture. This may result in the trainees taking compensatory actions before diagnosis is completed. Study of how these actions interact with diagnosis and how the model should be modified to incorporate the compensatory actions are potential

research problems.

Comparison of performance between naive and expert subjects will be an interesting extension to this research. This comparison can help understand the experts' strategy better which is important for building intelligent aiding systems. Also the differences in strategies used may indicate how people become experts.

As pointed out in Chapter VI, this research considered only the empirical constraints on modeling the experts' fault diagnosis performance. To make this model implementable as an intelligent diagnostic aiding system, much work needs to be done to explore the computational constraints.

Finally, research using expert subjects on realistic environments should be continued. A major difficulty lies in the gathering and analysis of protocols. Understanding the strategies used is difficult. More efficient methods should be developed to extract useful and systematic knowledge from the protocols or data automatically. Some preliminary research has been done by Waterman and Newell (1975), Bauer (1975), and Ericsson and Simon (1984). However, an automatic protocol analyzer has not been developed.

Appendix A

List of Failures Used in the Pilot Experiment

Failure No.	Faulty Component	Symptoms
1	STBD Economizer	STBD boiler high water alarm sounds. STBD boiler drum pressure decreases.
2	Condensate Pump	Main vacuum goes down as plant is brought to sea speed. Main vacuum alarm sounds.
3	FO Htr Stm Regulator	Smoke alarms & low fuel oil temp alarms sound. Black smoke is evident.
4	PORT FO Settling Tank	FO Header Pressure Low alarm sounds. STBD & PORT FO pressure low alarms sound.
5	PORT FO Settling Tank	All flame failure alarms sound, low fuel oil pressure alarm sounds. All fires are extinguished.
6	Gland Seal Regulator	Low vacuum alarm sounds on main condenser. Vacuum steadily dropping on slowing to half ahead during maneuvering.
7	Deaerating Feed Tank	High oxygen level in water in both boilers.
8	Supply Regulator	PORT boiler starts panting. PORT smoke alarm sounds.
9	Throttle Oil Booster Pump	Main engine RPM marginally above preset limit.
10	PORT Attenuator Ctrl Valve	High superheated steam temp from PORT boiler.
11	Condensate Pump Suc. Valve	Main condenser low vacuum alarm sounds. Main vacuum is steadily dropping.
12	FO Heater Steam Traps	Smoke alarms go off on both

		boilers.
13	Sea Strainer	Condenser vacuum low alarm sounds.
14	STBD Air Register 1	Repeated flarebacks on lighting off STBD boiler.
15	Fuel Oil Heaters	Low fuel oil temp alarm and smoke alarms on both boilers sound upon increasing speed.
16	Control Air Dryer	Losing main vacuum on slowing to half ahead from full ahead. Low vacuum alarm sounds.
17	FO Pump Suc. Strainers	Unable to maintain steam pressure & temp at increased speed.
18	STBD Boiler Drum Dry Pipe	Low rumbling noises from HP turbine at high RPM.

Appendix B

List of Failures Used in Sessions 1 to 5

***** Failure 1:

Initial Condition: The ship is under way at full speed (90 RPM).

Symptom: Condensate pressure low alarm sounds.

Failed Component: No. 100---Condensate pump
(broken flexible coupling)

Gauges Affected:

36---low
19---slightly dropping
27---low
58---low
55---high
60---high

Components Affected:

100---amperage low
104---in opening position
105---closed
78---condenser level high and overboard temp slightly high
79---hotwell full

***** Failure 2:

Initial Condition: The ship is under way at slow speed (20 RPM)

Symptom: When speeding up to full speed (70 rpm),
the No. 2 boiler level drops low.

Failed Component: No. 139---No. 2 feedwater regulator valve
(stuck open)

Gauges Affected:

58---slightly raise
57---low
56---level low
50---slightly high
69---15 psi
71---15 psi
73---7 psi
75---6 psi
77---5 psi
79---10 psi
81---3 psi

83---3 psi
85---5 psi
87---5 psi
89---9 psi
91---5 psi
93---9 psig

Components Affected:

104---closed
105---open
113---slightly open
139---open 10%

Indicators Affected:

188, 190, 192 ---- all extractions closed
256 ---- scoop closed
259, 231, 233, 235

***** Failure 3:

Initial Condition: The ship is underway at full speed (90 RPM).

Symptom: Main condenser vacuum low alarm sounds.

Failed Component: No. 84---Sea scoop (dirty)

Gauges Affected:

27---low
35---low
55---high
94---disc. pressure low
96---disc. temp. higher than normal

Components Affected:

78---Condenser shell is warm

***** Failure 4:

Initial Condition: The ship is underway at reduced speed (50 RPM).

Symptom: Main condenser vacuum low alarm sounds.

Failed Component:

No. 87---Sea strainer

Gauges Affected:

27---low
35---low
55---high
69---15 psi

71---15 psi
 73---8 psi
 75---7 psi
 77---5 psi
 79---7 psi
 81---3 psi
 83---3 psi
 85---8 psi
 87---7 psi
 89---9 psi
 91---7 psi
 93---9 psi
 94---pressure low
 96---temp higher than normal

Components Affected:

78---Condenser shell is warm and condenser level is low.
 84---scoop is closed.
 105---open 50%

Indicators Affected:

188, 190, 192 ---- all extractions are closed
 256 ---- scoop closed
 259 ---- low sea suction open
 231, 233, 235 ---- No. 1 cooling water circulator on
 also the discharge and suction valves open.

***** Failure 5:

Initial Condition: The ship is underway at full speed (90 RPM).

Symptom: Rumbling noise is heard from the boilers. The No 1 boiler starts to smoke and trips. Rumbling stops.

Failed Component:

No. 1 ---- No. 1 forced draft fan

Gauges Affected:

1---0 psig
 2---0 psig
 3---low
 4---low
 5---low
 62---black smoke

Components Affected:

1---amperage is low and motor is running
 33---closed
 34---closed
 31---closed

Indicators Affected:
138, 139

***** Failure 6:

Initial Condition: The ship is underway at full speed (90 RPM).

Symptom: High drum water level in no. 2 boiler sounds. The no. 2 boiler water level continues to rise.

Failed Component:
No. 139 ---- No. 2 feedwater regulator (broken diaphragm)

Gauges Affected:
58---slightly low
57---high
93---0 psig

Component Affected:
56 ---- Gauge glass reads high
104 ---- open position
139 ---- 100% open

***** Failure 7:

Initial Condition: The ship is underway at full speed (90 RPM).

Symptom: High pressure turbine vibration excessive alarm sounds and the engine noise level is higher.

Failed Component:
No. 52 ---- No. 2 internal attemperator

Gauges affected:
52---20 degree low

Components Affected:
76---- noise abnormal

***** Failure 8:

Initial Condition: The ship is underway at full speed (90 RPM).

Symptom: Vibration and noise and speed fluctuation occur in the no. 1 feedpump (on line). When the engineer starts the no. 2 pump, the same symptoms occurs immediately.

Failed Component:
No. 54 ---- no. 1 internal desuperheater

Gauges Affected:

51---20 degrees low

Components Affected:

125--- noisy

***** Failure 9:

Initial Condition: The ship is under way at full speed (90 RPM).

Symptom: There is a white haze in the no. 1 boiler periscope.

Failed Component:

No. 207 --- no. 1 boiler tubes

Gauges Affected:

1---high
2---high
3---high
4---high
5---high
6---slightly high
7---slightly low
19---slightly low
62---white smoke
44---slightly low
64---high
77---8 psig
89---4 psig

***** Failure 10:

Initial Condition: The ship is underway at full speed (90 RPM).

Symptom: No. 2 boiler trips. White smoke alarm, rumbling noise from boiler 2. FO pressure low alarm sounds.

Failed Component:

No. 30 --- FO Control Valve

Gauges Affected:

10---higher than normal
11---higher than normal
12---higher than normal
13---higher than normal
14---higher than normal
15---very low
16---low
17---low

18---low
52---slightly high
63---white smoke
65---low

***** Failure 11:

Initial Condition: The ship is underway at full speed (90 RPM).

Symptom: High superheated steam temperature in no. 2 boiler.

Failed Component:

No. 48 ---- no. 2 attemperator control valve

Gauges Affected:

52---high

Components Affected:

48----closed

***** Failure 12:

Initial Condition: The ship is underway at full speed (90 RPM).

Symptom: Fuel oil pressure low alarm sounds for both boilers.

Fuel oil pressure is erratic. White and black smoke
appear alternatively.

Failed Component:

No. 19 ---- Fuel oil settling tank (water)

Gauges Affected:

6---erratic

15---erratic

39---erratic

62---alternate between white and black smoke

63---alternate between white and black smoke

64---flow level fluctuating

65---flow level fluctuating

Components Affected:

21---noisy and discharge pressure erratic

***** Failure 13:

Initial Condition: The ship is underway at full speed (90 RPM).

Symptom: Low vacuum alarm sounds and vacuum is dropping when
slowing down the ship.

Failed Component:

NO. 168 ---- Gland seal regulator

Gauges Affected:

27---low
55---high
31---low

Components Affected:

168---low pressure output

******* Failure 14:**

Initial Condition: The ship is underway at full speed (90 RPM).

Symptom: FO temp low alarms sound on both boilers. Smoke alarms sound. Boilers indicating black smoke.

Failed Component:

No. 26 ---- fuel oil heater steam regulator

Gauges Affected:

41---low
47---low
62---black smoke
63---black smoke

******* Failure 15:**

Initial Condition: The ship is underway at full speed (90 RPM).

Symptom: When the bridge requests a speed reduction to 70 RPM, the no. 1 boiler periscope becomes hazy.

Failed Component:

No. 3 ---- no. 1 boiler air damper (stuck open)

Gauges Affected:

1---high
2---high
3---high
4---high
5---high
46---increasing
62---hazy(white smoke)
64---higher than normal

Components Affected:

3---open excessively

***** Failure 16:

Initial Condition: The ship is underway at full speed (90 RPM).

Symptom: Both boilers are smoking.

Failed Component:

No. 24 ---- fuel oil heater steam trap

Gauges Affected:

41---low

47---low

62---dark smoke

63---dark smoke

Components Affected:

26----fully open

***** Failure 17:

Initial Condition: The ship is underway at full speed (90 RPM).

Symptom: Steady loss of water in no. 2 boiler.
Boiler trips on loss of all flame.

Failed Component:

No. 208 ---- no. boiler tubes

Gauges Affected:

16---low

17---low

18---low

50---lower than normal

52---slightly low

57---low

58---low

Components Affected:

56----gauge glass level dropping

104---open

139---fully open

***** Failure 19:

Initial Condition: The ship is underway at reduced speed (50 RPM).

Symptom: A drop in vacuum occurs after speed was reduced.

Failed Component:

No. 157 ---- 35 system dump regulator

Gauges Affected:

27---low
23---lower than normal
55---higher than normal

Components Affected:

78---overboard temperature slightly high
89---fully open
157---open 40 %

Indicators Affected:

188, 190, 192 --- all extractions closed
256 --- scoop closed
231, 233, 235 --- 1 circulator is running with valves open

***** Failure 20:

Initial Condition: The ship just came out of shipyard and
is underway at full speed (90 RPM).

Symptom: Ocassionally experiences sudden drop and rise of
superheater outlet pressure and a sudden drop
and rise of turbine speed.

Failed Component:

No. 49 ---- 1 superheater (blockage)

Gauges Affected:

8---suddenly rise and drop
1---suddenly rise and drop
2---suddenly rise and drop
3---suddenly rise and drop
4---suddenly rise and drop
6---suddenly rise and drop
7---suddenly rise and drop
9---suddenly rise and drop
46---suddenly rise and drop

***** Failure 21:

Initial Condition: The ship is underway at reduced speed (50 RPM).

Symptom: The command is given to go up to 90 RPM. The ship goes to
90 RPM. When tank cleaning is started the boiler will not
maintain pressure without excessive black smoke in PORT
boiler.

Failed Component:

NO. 11 ---- 2 air heater (dirty)

Gauges Affected:

10---high
11---high
12---high
13---high
14---low
48---slightly low
63---dark smoke

***** Failure 22:

Initial Condition: The ship is underway at full speed (90 RPM).

Symptom: When testing the boiler water, 1 boiler has an abnormal drop in alkalinity, sulfite and phosphates with no change in chloride content. In the 2 boiler the chemical concentration increased.

Failed Component:

No. 53 --- internal desuperheater (small leak)

Gauges Affected:

45---slightly low

***** Failure 23:

Initial Condition: The ship is underway at reduced speed (70 RPM).

Symptom: When raising to normal full speed, boiler pressures drop and smoking in both boilers. Low atomizing steam alarms sound.

Failed Component:

No 90 ---- atomizing steam reducing station

Gauges Affected:

6----fluctuating
15---fluctuating
7----low
8----low
9----low
16---low
17---low
18---low
28---low
40---low
62---dark smoke

63---dark smoke
 69---15 psi
 71---15 psi
 73---8 psi
 75---7 psi
 77---5 psi
 79---7 psi
 81---3 psi
 83---3 psi
 85---8 psi
 87---7 psi
 89---9 psi
 91---7 psi
 93---9 psi

Components Affected:
 90---open 100 %

Indicators Affected:
 188, 190, 192---all extractions are closed
 256---scoop closed
 259---low sea suction open
 231, 233, 235---No. 1 cooling water circulator on also
 the disc. and suction valves open.

***** Failure 24:

Initial Condition: The ship is underway at full speed (90 RPM).

Symptom: Condensate pump discharge pressure lower than normal.
 Low Condensate P. alarm sounds.

Failed Component:
 No. 100 ---- condensate pump
 (pump is running and has an internal failure)

Gauges Affected:
 19---slightly low
 27---slightly low
 36---low
 58---low and falling
 60---high

Components Affected:
 79---high level and rising
 100---low amperage
 104---open

***** Failure 25:

Initial Condition: The ship is in port at dockside

Symptom: The low control air pressure alarm sounds on the main console.

Failed Component:

No. 175 ---- control air compressor

Gauges Affected:

38---low and falling

Components Affected:

175---electric motor running and broken drive belt

***** Failure 26:

Initial Condition: The ship is underway at full speed (90 RPM).

Symptom: Deaerating feed tank level is marginally low
and slowly decreasing.

Failed Component:

No. 116 ---- Atmospheric drain tank level transmitter

Gauges Affected:

58---marginally low

61---very high

Components Affected:

104---slightly open

115---tank is overflowing

***** Failure 27:

Initial Condition: The ship has just received departure order
and is steaming at full speed (90 RPM).

Symptom: Marginal loss of main condenser vacuum.

Failed Component:

No. 105 ---- Condensate recirculation regulator

Gauges Affected:

27---marginally low

58---slightly low

60---marginally high

Components Affected:

79---high level

105---open 80 %

***** Failure 29:

Initial Condition: The ship is underway at full speed (90 RPM).

Symptom: Marginally low saltwater service system pressure.

Failed Component:

No. 171 ---- Saltwater service pump strainer

Gauges Affected:

34---marginally low

53---marginally high

Components Affected:

72---marginally high lube oil temperature

Appendix C

Protocols

Failure 19 (Vacuum drop at reduced speed) (S31)

- L100: First indication is a lot of vacuum loss.
L101: Lo, first thing I'll check will be the gland seal
L102: (to see) if it has pressure after regulator.
L103: You know on the board it is at least, say, 1 or 2 pounds.
L104: If that is OK, I believe, you check circulator maybe,
L105: to see that you have enough water going back into the condenser
L106: to keep the air ejector cool.
E107: Why the air ejector?
L108: Oh! If you talk about the vacuum pump... I am not quite familiar with the vacuum pump. Let us say the
L109: I know with the air ejector you lose vacuum that way.
L110: What else can be wrong?
E111: Lo, those are most common problems?
L112: Yes, I would think so, unless you...
L113: Assuming the condensate pump running normally,
L114: that will be the first indication, the major cause.
E115: Lo, what else?
L116: Check level of condenser.
E117: What do you want to find?
L118: You want to find out if it is normal,
L119: if it is too high then... you see it could be getting water from somewhere else.
E120: What do you mean by somewhere else?
L121: Say, for example, condensate pump may not be functioning properly.
E122: Think aloud.
L123: It would be ..., well. If the level is OK, then you look for something else.
L124: We just slow down.
L125: You assume everything else is OK.
L126: In other words, circulating water is fine,
L127: no reason for that, which you think probably will be ...
L128: I don't think it has any problem with the bleeds
L129: I don't know. The symptom is...
E130: Vacuum dropping when slowing down to 50 from full speed.
L131: Possibly the bleeds, you could lose it that way.
E132: Any other possibilities?
L133: Like I said, the dump could cause it,
L134: when you have problems with the 35 lb line.
L135: If that was the problem you hear the 35 lb make up and you hear the feedpump speed up a little.
L136: You definitely hear the 35 lb line making up
L137: from the high pressure steam coming into the low pressure steam
L138: where the low pressure dumping it into the condenser.

- L139: So, one failure would cause the reaction on the make up and you hear that
- L140: you could remedy that problem right away.
- E141: Even before the vacuum drop?
- L142: I would think yes.
- L143: If that was the case, vacuum would be secondary
- L144: because you would have other problems. On our ship we have other things.
- L145: You could... You have deck return coming into the main unit. But that should not be...
- L146: Say at full speed
- L147: it would not be making any effect. I don't believe from other....
- L148: On a vacuum, you basically looking for a leak
- L149: getting into the main unit somehow.
- L150: The turbine... that is basically the problem.
- E151: How about the deaerator? If there is a hole in the deaerator.
- L151: What type of leak?
- L152: There is a pressure on the deaerator.
- L153: Deaerator is under 35 lb pressure aha... goes into the feedpump
- L154: so I don't think there is any problem there.
- L155: If the level of the deaerator changes,
- L156: you'll notice if the makeup for the dump fail for the deaerator,
- L157: and you'll notice the level change.
- L158: I don't believe it will affect the vacuum at all.
- E159: And surely the vacuum pump itself can fail that will cause the vacuum to drop, but it should not happen when slowing down.

Failure 23 (When raising to normal full speed, boiler pressures drop and smoking in both boilers. Low atomizing steam alarms sound.) (S32)

- L200: First we need to find out what's causing the boiler pressure to drop.
- L201: Check fuel oil pressure. You said it is OK.
- L202: And you said smoking.
- L203: That'll probably indicate the pressure would be extremely high
- L204: because it is a large load from the boiler.
- L205: So that could cause the smoke.
- L206: It would be too much demand for the boiler to make up.
- L207: That's the reason the boiler pressure is dropping.
- L208: With low atomizing steam we have to check the 150 lb system to see if that is normal.
- L209: If that's normal, we have to go to the atomizing steam regulator to see that is functioning properly.
- L210: If that fails, you have poor atomization which undoubtedly will trip the boiler.
- E211: Cause the smoke?
- L212: Surely, because you have low atomizing steam,
- L213: and you would have to slow down to bring the pressure back so you don't lose the plant.
- L214: Doing that will bring up the fuel oil pressure back.
- L215: Remedy the atomizing steam somewhat or bypass it and find out the fault. That seems to be the problem.
- L216: What if the atomizing steam pressure alarm doesn't go off.
- L217: Need to find out whether there is too much demand for the boiler,
- L218: whether the boiler pressure is dropping in the auxiliary system or the main system.
- L219: In order to bring the pressure back, you have to reduce the load.
- L220: We'll have to come back to a reduced speed
- L221: for sure to secure some other steam demand and find out the problem.
- E222: What's the procedure to find out the problem?
- L223: You have to check to see if both burners in both boilers are operating properly.
- L224: If the fuel oil is following the demand and the air,
- L225: if there is not enough air we won't get proper atomization. I mean proper burning. It could be.
- L226: Have to check the different pressure.
- L227: Whether the temperature of the boiler is OK
- L228: to see if it would have dirty tubes.
- E229: Blockage?
- L230: Anything could be blocked.
- L231: It could be having extremely cold water going into the boiler.
- L232: So the economizer may be the problem?
- L233: Possibly could be economizer plugged up, or...
- E234: Remember we have smoke. Does it help to narrow down the possibilities?
- L235: Supposing the atomizing steam is proper?
- E236: Yes.

- L237: Not enough air.
L238: Because if it is not enough air, fuel oil temperature is too low,
L239: it will be smoke.
E240: You assume black smoke?
L241: I like to say it is too much demand whether it is enough heat transfer for the fire to the burner.
L242: Or fuel oil pressure too high then you have smoke because too much demand.
E243: Too much supply?
L244: Yes, oh... I guess, oh... too much demand on the boiler.
L245: It will try to make up more fuel.
L246: More fuel will.... so you exceed the air,
L247: not enough to burn them, so you see it black.
E248: If only boiler pressure drops and black smoke are noticed, What is the most possible reason?
L249: Not enough air I think.
E250: How does the air affect the boiler pressure?
L251: Boiler pressure? Well, if the fuel follows the air it will drop.
L252: If the fuel is following the air, you shouldn't have black smoke. However, we have black smoke here.
L253: The combustion control, the air damper must not be open right.
L254: The speed of the forced draft fan might not be up speeded.
L255: Or... I mean it is a process of elimination.
L256: I mean it is hard to sit here and tell you what's wrong.
E257: But you can assume, see, suppose the combustion control is not functioning well, the ratio is not right.
L258: You don't have enough air.
L259: You not going to burn right.
L260: And you won't have the right amount of mixture for the air and fuel to give you proper burning.
L261: So you going to lose pressure.
E262: Does it happen on any condition?
L263: It really doesn't matter.
L264: If you can't get the right fire, you not going to get pressure.
L265: I mean unless there is no demand. You going to get the pressure drop.

How does the power plant work? (S31)

- L301: A boiler would generate the steam,
- L302: the steam leave the boiler at a predetermined pressure and temperature.
- L303: It leaves the boiler at different levels at different outlets,
- L304: e.g. superheated steam, desuperheated steam, each for different purposes.
- L305: The superheated steam goes into the throttles for the main turbine.
- L306: That adjusts the direction or the amount of steam goes it turbines.
- L307: Turbines use the heat and steam to get the work out.
- L308: And after it is used up the thermo energy,
- L309: it is cool and changed back from steam to water back in the condenser.
- L310: The condensate is pumped from within the air ejector up into the dc heater.
- L311: Dc heater will heat it and the deaerator take the air out of it.
- L312: You don't want the air to enter the boiler system.
- L313: From there it goes into the feedpump.
- L314: And return to the boiler and heated in different stages.
- L315: So it is a closed loop per se.
- L316: Then you have all the auxiliaries which operate out of different pressure steam to maintain the powerplant.
- L317: We have 35 lb line which is used for gland seal.
- L318: For the deaerator you have the 150 lb line which is used for the make up.
- L319: 150 lb line for the fuel oil heater
- L320: and regular steam which is desuperheated steam (600 lb) is used for the feedpump.

How does gland seal work? (S31)

- L401: What gland seal is trying to do is to prevent air from leaking
L402: from outside of the casing into the casing eventually dump into the vacuum.
L403: What it does is to seal off the air which would be sucked in otherwise
L404: because there is a difference in pressure between the vacuum and the atmosphere.
L405: So under full speed with steam admitting into the turbine at a high enough pressure
L406: when the steam is trying to escape, you don't need gland seal because there is pressure within the casing.
L407: So what you do is keep that steam in that casing, so air can't come in.
L408: So the gland seal prevents the air you don't want in the system.
L409: In reduced speed you would need the gland seal.
L410: You have to be able to maintain the vacuum.
L411: It is created with the steam changing back.
L412: The volume of steam going back to volume of water which creates the vacuum.
L413: Also it keeps the temp, e.g. if you have air sucking in on the rotor, it is not a good condition.
L414: It may have a tendency to bolt the rotor or something to that effect if it is in a standing position.
E415: Do you have any experience when you were on a ship, some problem happened that reminded you of some other symptoms.
L416: You don't always remember what happened.
L417: I may not remember the exact cause of the effect. But I may be able to....
L418: just by experience in my own mind that I know to check this one thing,
L419: a direction, which I believe is the definition of experience.
L420: Say, if I'm at full speed, and vacuum is coming down.
L421: I'll look the level of the condenser because 7, 10 years ago I was a third, we got a departure,
L422: this happened to me, I didn't know what to do at that time.
L423: But down the line, this did happen to me again.
L424: But now it is such a common thing to me. Of course, the experience I never forget.

Failure 8 (Vibration, noise and speed fluctuation occurred in the no. 1 feedpump. When the engineer started the no. 2 pump, the same symptoms occurred immediately.) (S34)

L450: Well, maybe loss of suction because that is common to both pumps.
L451: Check the level of the deaerator (normal).
L452: I have checked out the suction side of the water side of both pumps.
L453: Now I will check out discharge side of both pumps.
L454: Check out the condition of the condensate pump system. This will give me what valves are open.
L455: OK, next one on the water side, HP heater (normal).
L456: What will cause fluctuation. Maybe there is
L457: I will go to my steam system.
L458: Now I have eliminate the suction water side of the pump. I am up to the economizer.
L459: Check temp gauges coming in and out of economizer.
L460: I wonder if there is a leakage in the economizer?
L461: Maybe the economizer is leaking.
L462: Sometimes when you started up a pump it will cause speed fluctuation and trip out.
L463: Inlet is normal; check the outlet.
L464: If outlet is low, then that means the economizer is plugged up.
L465: Normal. So every thing on the water side is OK.
L466: Now I am going to the steam side of the feedpump.
L467: Next thing I am looking for is what is common to both pumps.
L468: Check if 35 lb line has back pressure....
L469: Check if 35 lb pressure is normal (normal).
L470: If that is normal, back pressure cannot be low.
L471: I will be getting.... Pump would not be fluctuated.
L472: Now go on to the steam side of the inlet.
L473: I cannot be nothing that is common...
L474: I am going to be backtracking.
L475: Desuperheated steam for my feedpump....
L476: Check if both of them come out right.
L477: I want to find a cause common to both feedpumps.
L478: I have checked my pressure and levels.
L479: I wonder about if the temp of the steam is relevant...
L480: If cold water goes into the feedpump,
L481: It would cause it to flash the steam
L482: And the pumps would lose suction.
L483: A lot of times when we have problems with our deaerator
L484: And we have to bypass the deaerator.
L485: We take water from our distilled tank.
L486: The water in the deaerator is 275 degrees,
L487: And the water coming from the distilled tank is 90 degrees.
L488: Difference in change will cause it to flash.
L489: And the pump will lose suction.
L490: I am going to check out the temp of the water coming from the deaerator.

L491: I am going to condensate system again.
L492: Sometimes it is lube oil, I may lose lube oil.
L493: Nothing there, what else can cause it?
L494: Maybe the lube oil cooler (pressure of salt water).
L495: It says OK.
L496: The only thing I can go on is my condensate system and steam system.
L497: We have eliminated condensate system because water is at normal pressure, tem and level.
L498: The economizer says OK.
L499: Steam temp may be a problem.
L500: 20 degrees low coming out of the desuperheater.
L501: Temp of steam could cause fluctuation of the feedpump
L502: Because the higher the temp the more the work you get out of the steam.
L503: So if it is 20 degrees low, it got to be something wrong in my desuperheater.
L504: I am now backtracking.... coming out of the superheater
L505: Going into the desuperheater.
L506: Backtrack a little bit further.
L507: Come out of the drum going to my superheater.
L508: I am going to check my drum level (normal).
L509: So, back into my superheater.
L510: It has to be the attemperator, the desuperheater or the superheater.
(At this point S33 found the solution ---- desuperheater failure)
E511: Why the pressure is normal but the temp is abnormal?
L512: Maybe there is a leak in the desuperheater.
L513: That causes loss of chemicals.
L514: Temp will be lower.
L515: If I have a leak it will go into my drum.
L516: My flow from the desuperheater outlet will be slower,
L517: And it would be cooling down too much.

Failure 9 (White haze is seen on STBD boiler periscope) (S33)



















- L550: White haze means too much air.
L551: Let us check the STBD burners' management system. It says it is running.
L552: Look at the air system. Too much air may be the problem.
L553: So burners' management system is OK. Let us start with the fan and work our way down.
L554: I got pressure to the boiler out of the fan.
L555: Signals go into the combustion control of that fan,
L556: that regulates the fan.
L557: Maybe the combustion control is sending a wrong signal to the top of the diaphragm.
L558: Or the diaphragm is broken.
L559: If it fails open then I have more air than I need.
L560: So check the air pressure.
L561: It is high.
L562: The fan is OK since it is running. Maybe the crossover..... normal.
L563: Go down the line see if I can pick up something else.
L564: Check windbox down the line. It says high..... but crossover is normal.
L565: So nothing is wrong along this line. FD fan is running.
L566: Maybe it has to be slow. Well, OK.
L567: I will say maybe my fuel oil has problem.
L568: My damper is staying normal but something happens to my fuel oil.
L569: Something has to stick.
L570: Crossover is OK, inlet damper is alright, outlet is lower all the way down my line.
L571: Unless I have a dirty burner....
L572: But that will give out black smoke.
L573: So I have eliminated my burner. My burner is OK.
L574: Let us check out the fuel.
L575: Let us check out the temp of the fuel.
L576: Temp is normal, pressure is slightly high.
L577: If that is high.... my air is high too.
L578: One supposes to follow the other.
L579: That is a part of the burners' management.
L580: Maybe something stuck along the fuel oil line.
L581: Fuel oil may then be slightly high.
L582: Combustion control sends the signal to my air system.
L583: So, let us check the regulating valve, check the flow before it.
L584: Slightly high.... I am going further up the fuel oil system.
L585: Fuel oil temp is OK, pressure is high.
L586: Check the strainer... no, that cannot be it.
L587: Because that should cause pressure to go low.
L588: Go back to the fuel oil pipe.
L589: Check the fuel oil pressure regulator.
L590: Normal.... Go back further.
L591: Why the pressure is high?

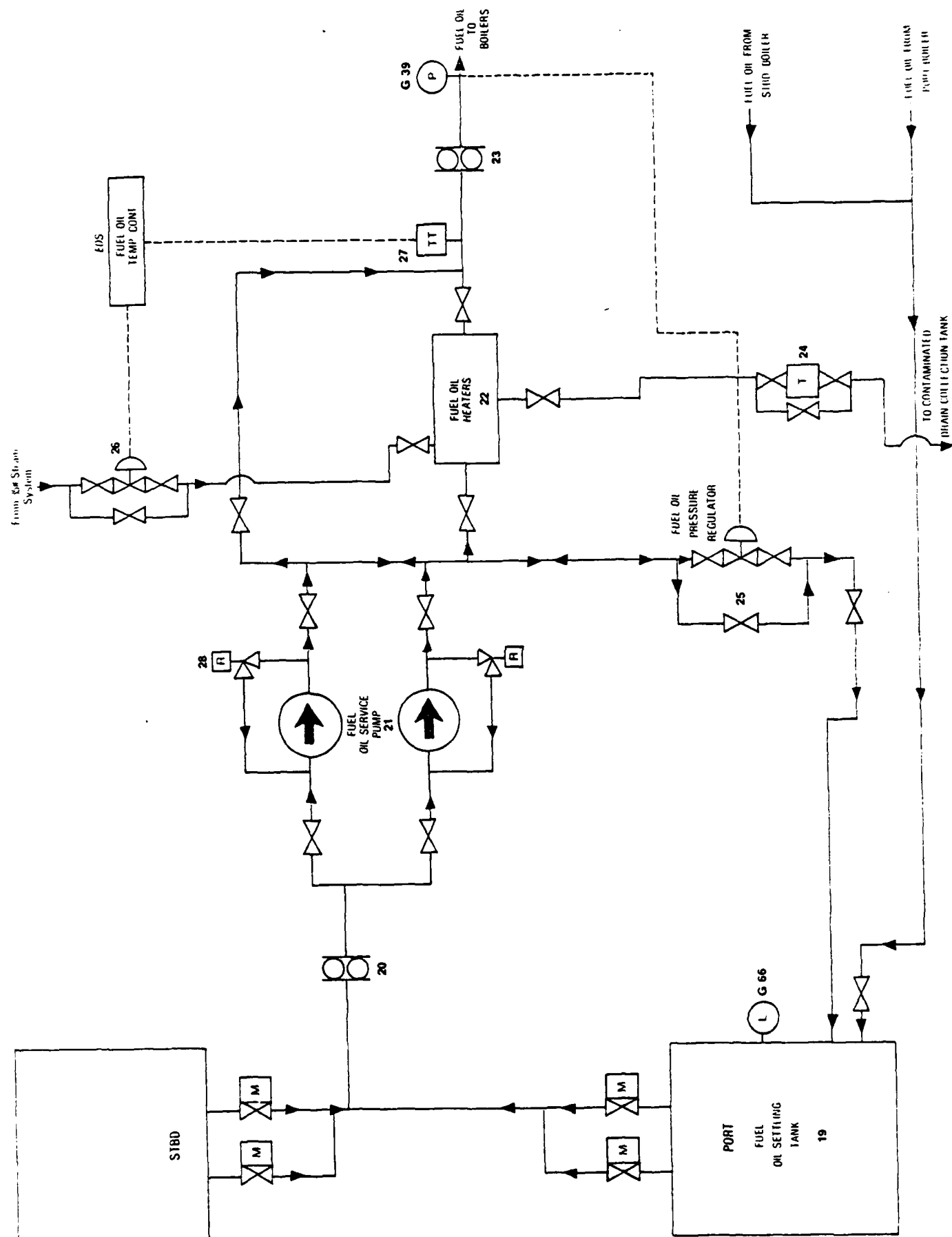
L592: Something wrong in the pump? No, it cannot be because it would be low too.
L593: If the relief valve is open....
L594: It would be low too.
L595: Go back to my burner again.
L596: Maybe my burner is clogged.
L597: But that would be black smoke.
L598: So..... temp is normal.....
L599: I am just going to say maybe it runs out another pump.
L600: I will check out the fuel oil pump.
L601: Pump is OK.
L602: Maybe my damper open too much.
L603: Check the other boiler see if pressure going in is high.... normal.
L604:slightly high fuel oil pressure and too much air....
L605: I am stuck.

Appendix D
System Schematics

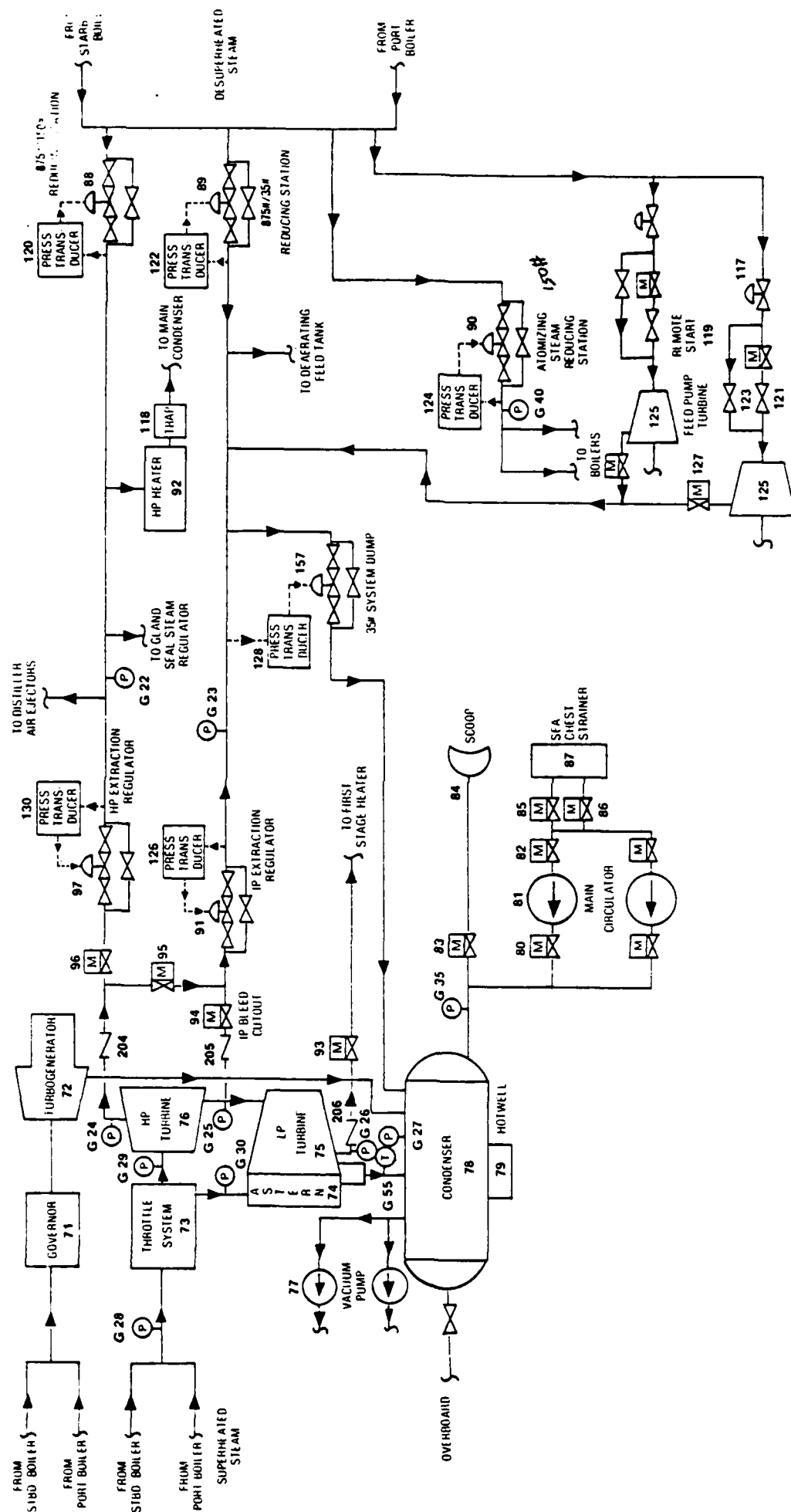
SYMBOLS

141

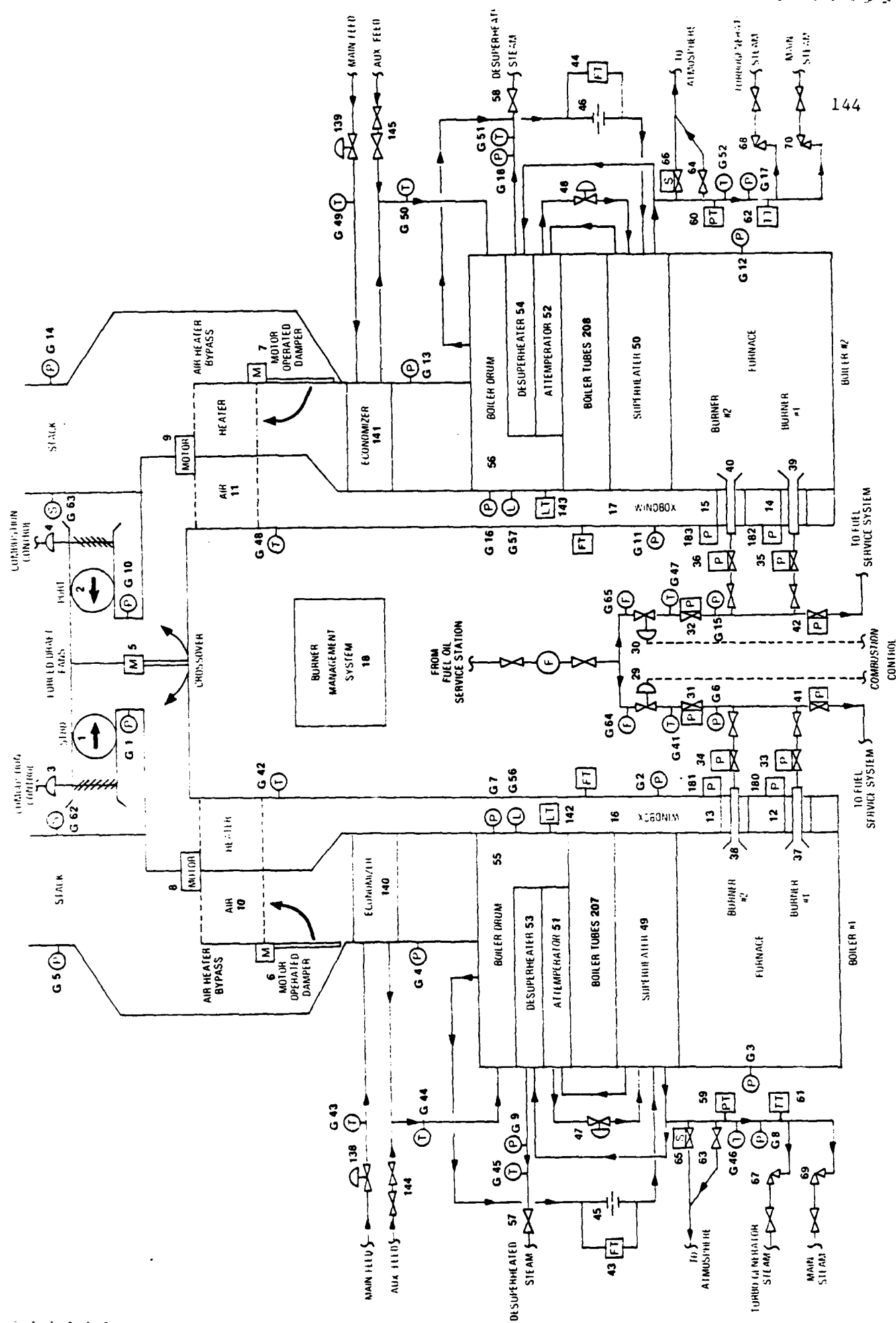
	PRESSURE INDICATOR
	TEMPERATURE INDICATOR
	LEVEL INDICATOR
	FLOW INDICATOR
	SMOKE INDICATOR
	VALVE
	ANGLE VALVE
	STRAINER
	PUMP
	LEVEL TRANSMITTER
	PRESSURE TRANSMITTER
	TEMPERATURE TRANSMITTER
	FLOW TRANSMITTER
	MOTOR OPERATED VALVE
	SAFETY VALVE
	RELIEF VALVE
	AIR OPERATED VALVE
	AIR ACTUATOR



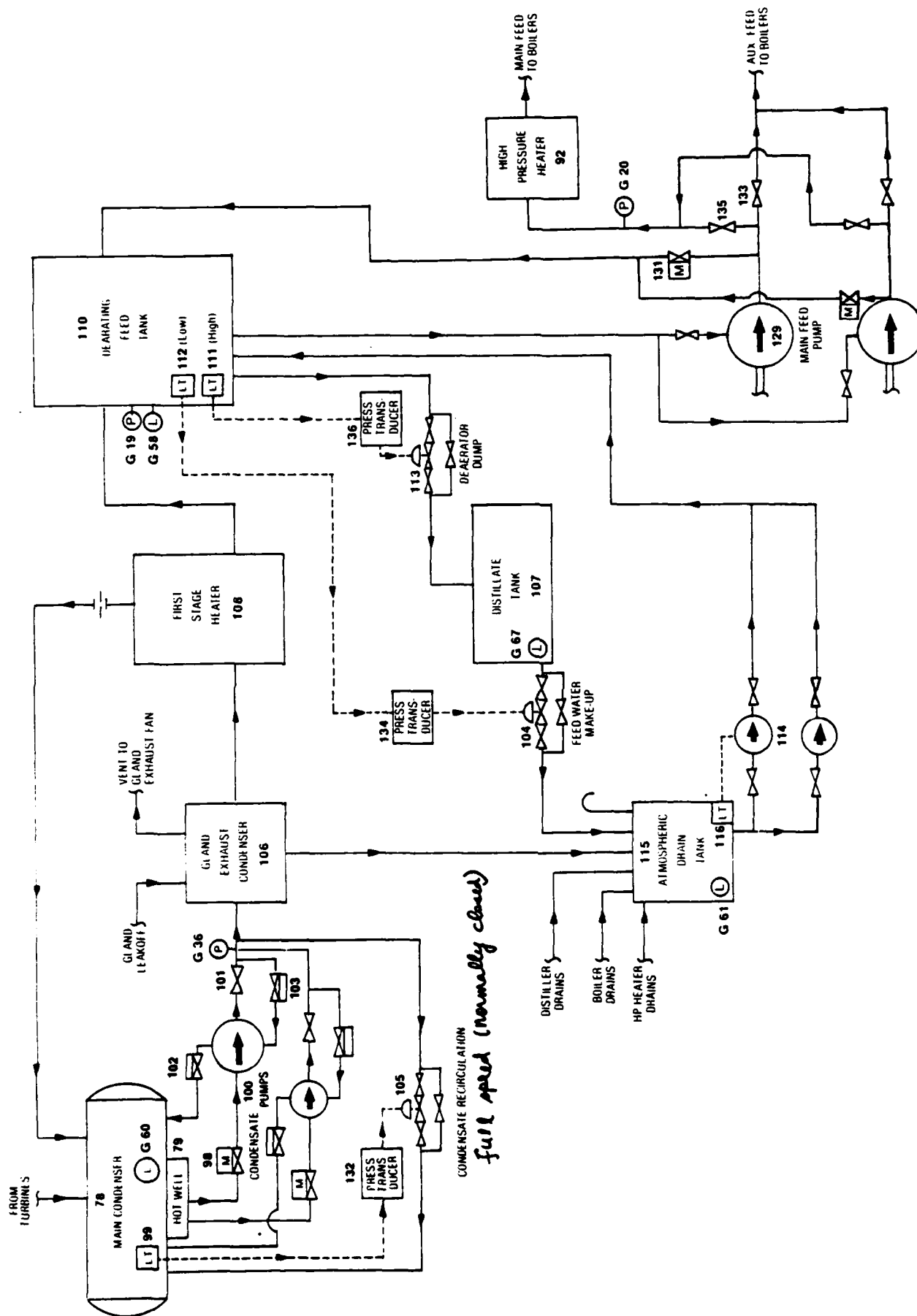
FUEL OIL SUPPLY SYSTEM



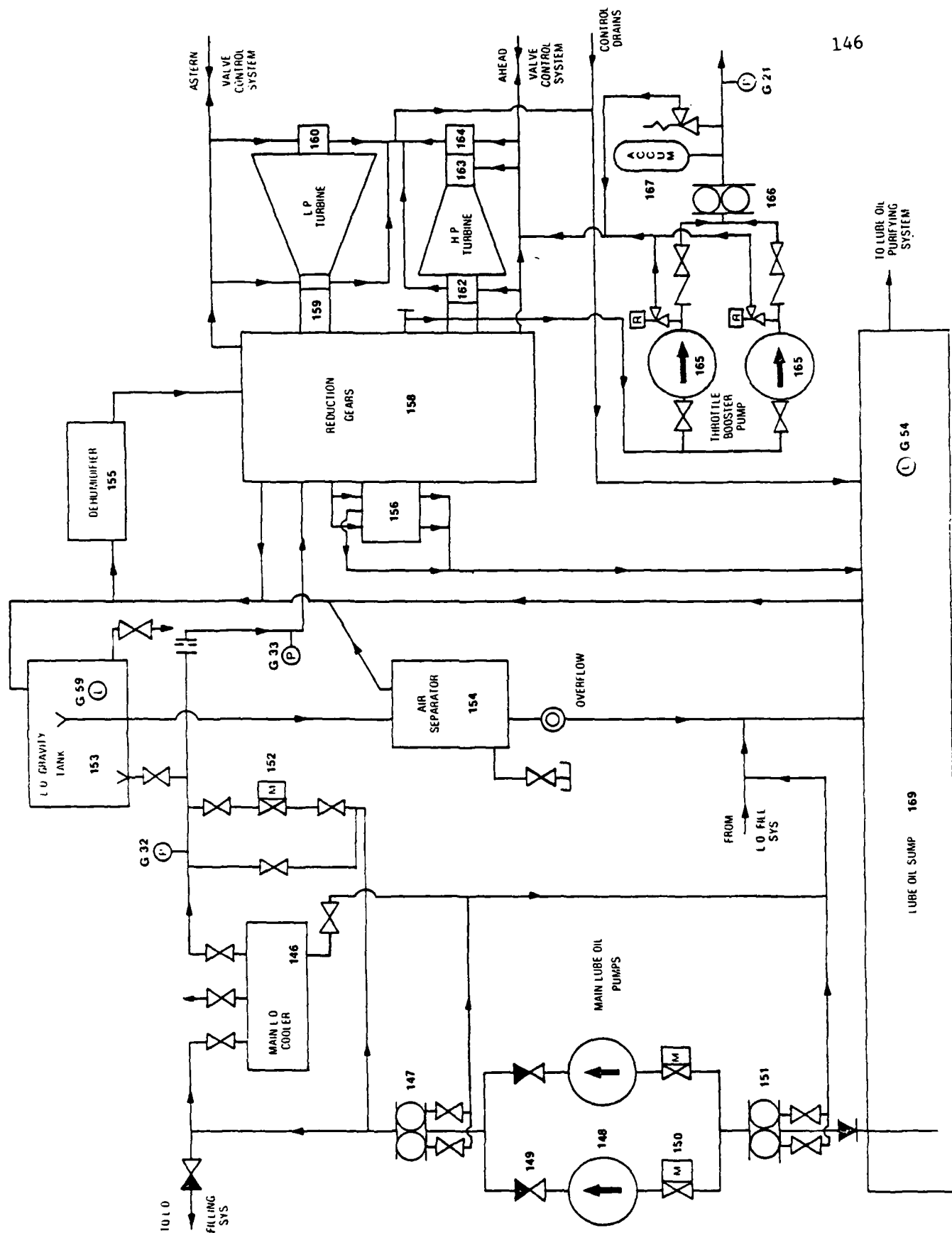
MAIN STEAM, AUXILIARY STEAM, AND CIRCULATING WATER SYSTEMS



BOILER SYSTEM

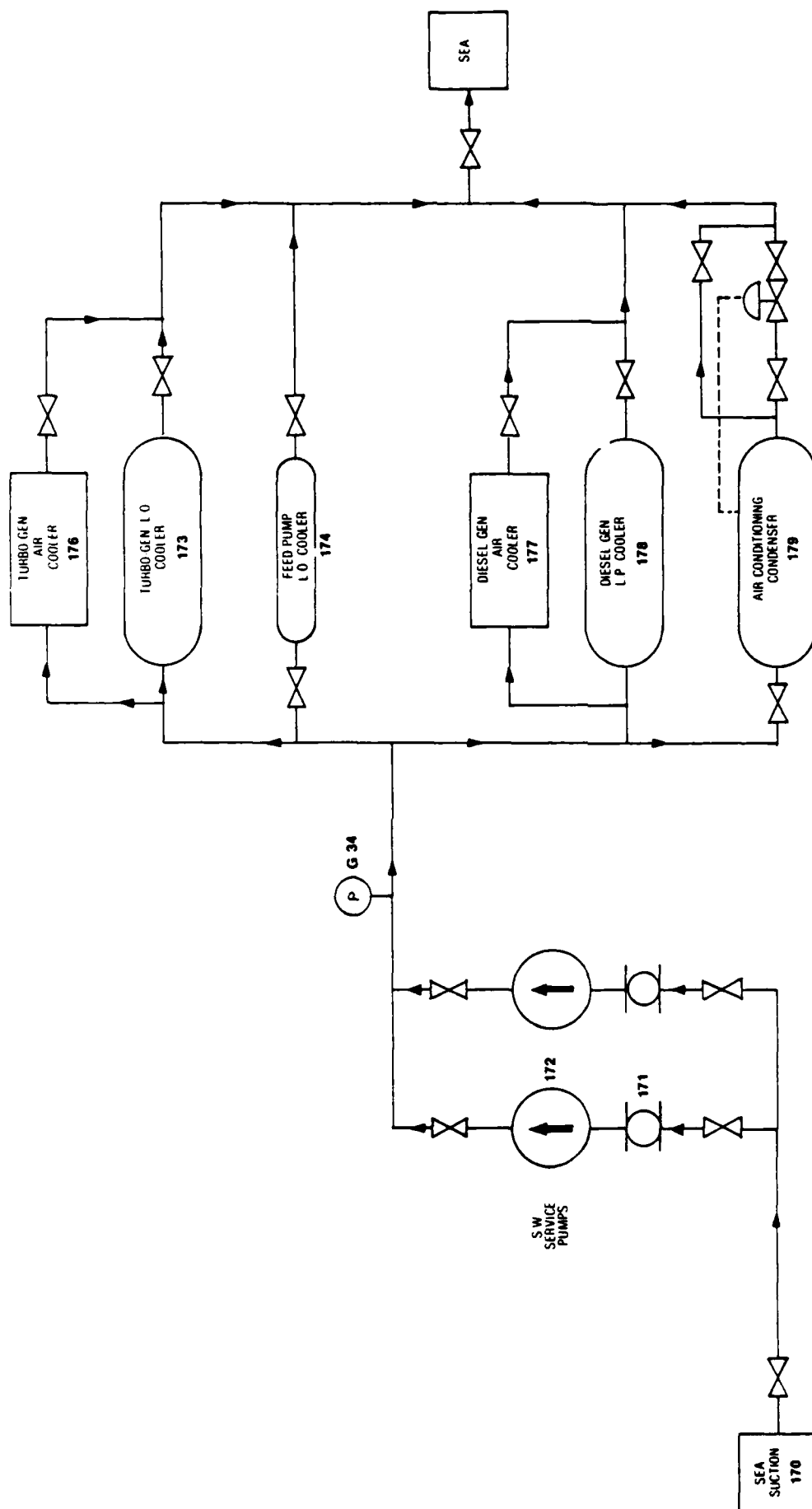


CONDENSATE AND FEED SYSTEM



146

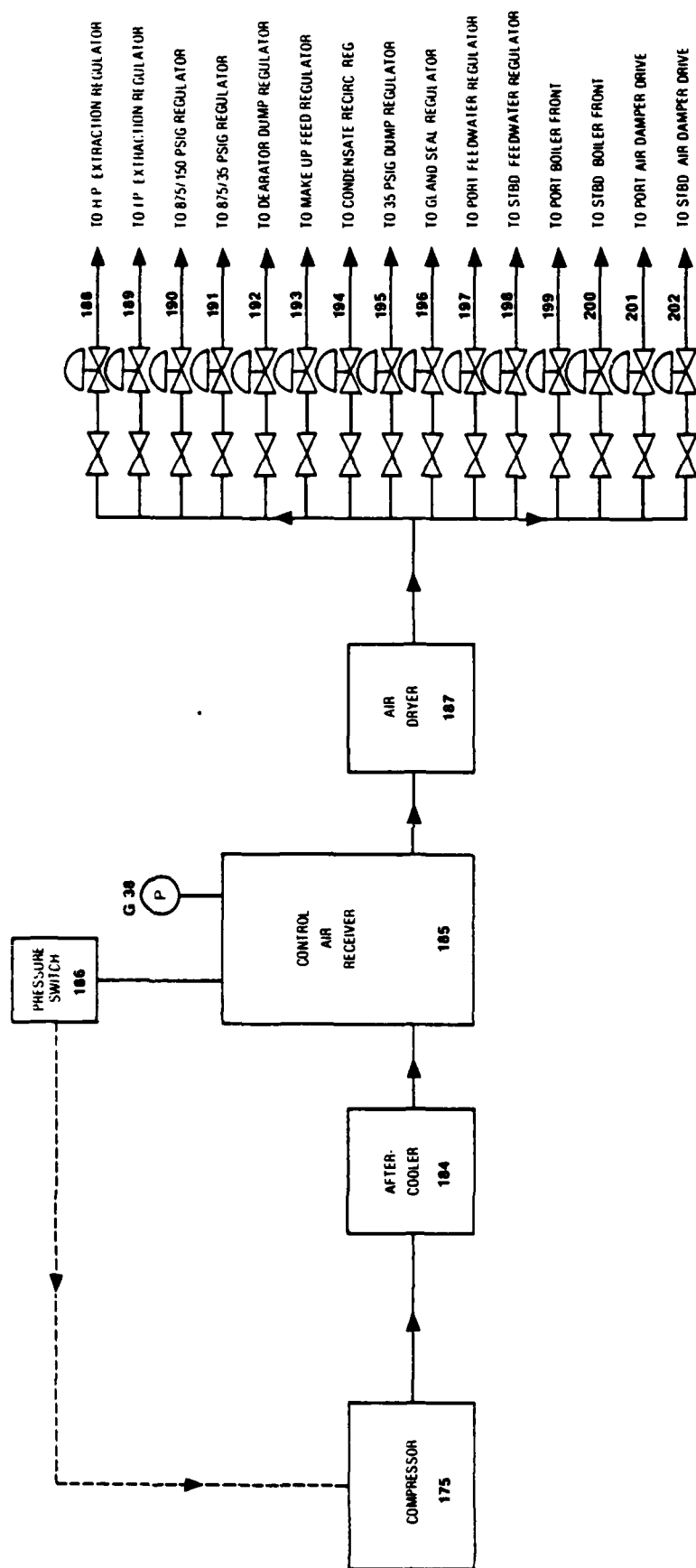
LUBRICATING OIL SYSTEM



SALTWATER SERVICE SYSTEM



GLAND SEAL SYSTEM



CONTROL AIR SYSTEM

Appendix E
Instruction Sheet

Marine Powerplant Troubleshooting

You are about to participate in an exercise where a marine powerplant is simulated on an Apple personal computer. The computer will display symptoms resulting from a realistic failure in a component or a subsystem of a marine powerplant. The symptoms correspond to the failure of a single component or a subsystem, or a single identifiable cause such as contaminated water. In addition to the computer display, you are provided with a set of schematic diagrams and a list of gauges and other status indicators. The schematic diagrams, gauges, and status indicators correspond to the engine room simulator at MSI.

You can check the status of the gauges and displays by interrogating the computer using a menu of commands. These commands enable you to check individual components, group of annunciators, diagnose and declare the failed component, or quit. A number of trial runs will be used to familiarize you with the simulation.

Each run incorporates a single failure. Status of relevant displays is stored in the computer for each failure. To check the displays corresponding to a subsystem, you must identify that system by its unique identification number. These numbers are indicated in the appropriate locations on the schematic diagrams.

Even though a failed component may result in continuous change in status in real life systems, the status of displays in our simulator has been frozen at a certain arbitrary point in time after the failure has occurred. Hence, you get a "snapshot" view of displays after the failure has occurred.

You may check as many gauges and displays as you want before identifying the failed component. However, you should try to keep the number of checks and readings to a minimum since in real life each reading would incur some cost. You are expected to identify the cause of the failure within the allotted time. Preferably, you should identify the cause as quickly as possible.

A number of training runs will be provided so that the exercise procedure becomes clear. By the end of these trials, you will be familiar with the command options on the computer, and the schematic diagrams. Please feel free to ask questions during any part of the trial runs.

Please do not discuss this exercise with any other trainee.

Appendix F
Command List

<u>COMMAND</u>		<u>GAUGE</u>
		<u>Boiler Gauges</u>
STBD.	PORT	<u>Pressures</u>
G1	G10	Forced Draft Fan Discharge Pressure
G2	G11	Wind Box Pressure
G3	G12	Furnace Pressure
G4	G13	Economizer Inlet Pressure (Gas Side)
G5	G14	Uptake Pressure
G6	G15	Fuel Oil to Burners Pressure
G7	G16	Boiler Drum Pressure
G8	G17	Superheated Steam Pressure
G9	G18	Desuperheated Steam Pressure
		<u>Temperatures</u>
G41	G47	Fuel Oil Temperature
G42	G48	Air Heater Outlet Temperature
G43	G49	Economizer Inlet Temperature
G44	G50	Economizer Outlet Temperature
G45	G51	Desuperheated Steam Temperature
G46	G52	Superheated Steam Temperature
		<u>Levels</u>
G56	G57	Boiler Drum Level
G62	G63	Smoke Indicators
		<u>Flow</u>
G64	G65	Fuel Oil Flow
<u>BOTH</u>		
G38		Control Air Pressure
G39		Fuel Supply Header Pressure
G40		Atomizing Steam Pressure
G66		Fuel Oil Settling Tank Level
		<u>Transducer and Transmitter Gauges</u>

SUPPLY LOADING

G68	G69	H.P. Extraction Regulator
G70	G71	I.P. Extraction Regulator
G72	G73	875/150 PSIG Reducing Station
G74	G75	875/35 PSIG Reducing Station
G76	G77	Make-up Feed Regulator
G78	G79	Condensate Recirculator Regulator
G80	G81	Deaerating Feed Tank Dump Regulator
G82	G83	35 P.I.C. System Dump Regulator
G84	G85	Atomizing Steam Regulator
G86	G87	Stbd. Boiler Steam Flow Transmitter
G88	G89	Stbd. Boiler Drum Level Transmitter
G90	G91	Port Boiler Steam Flow Transmitter
G92	G93	Port Boiler Drum Level Transmitter

<u>COMAND</u>	<u>GAUGE</u>
<u>Other Engine Room Pressures</u>	
G19	Deaerator Pressure
G20	Feedwater Header Pressure
G21	Throttle Hydraulic Oil Pressure
G22	150 PSIG System Pressure
G23	35 PSIG System Pressure
G24	H.P. Extraction Pressure
G25	I.P. Extraction Pressure
G26	L.P. Extraction Pressure
G27	Condenser Vacuum
G28	Throttle Supply Steam Pressure
G29	Ahead Turbine Pressure
G30	Astern Turbine Pressure
G31	Gland Seal Steam Pressure
G32	Lube Oil Header Pressure
G33	Main Engine Lube Oil Pressure
G34	Salt Water Service System Pressure
G35	Main Cooling Water Pressure
G36	Condensate Pump Pressure
G37	Turbo Generator Lube Oil Pressure
G94	Main Cooling Water Disch. Press.
<u>Other Engine Room Temperatures</u>	
G53	Feedpump Lube Oil
G55	L.P. Turbine Exhaust
G95	Main Cooling Water
G96	Main Cooling Water Discharge
<u>Other Engine Room Levels</u>	
G54	Lube Oil Sump Tank Level
G58	Deaerator Tank Level
G59	Lube Oil Gravity Tank Level
G60	Main Condenser Level
G61	Atmospheric Drain Tank Level
G67	Distillate Tank Level

COMMAND STATUS INDICATOR PAGES

I1	Throttle Control
I2	Main Turbine Bearing Temps
I3	Stbd. Boiler Burner Mgm.
I4	Port Boiler Burner Mgm.
I5	Salinity
I6	SSTG Bearing Temps
I7	Bleed Status
I8	Lube Oil System
I9	Condensor System
I10	Turbine Vixax Panel
I11	Fuel Oil System
I12	Feedpump Status

REFERENCES :

- Adams, J. A., "On the Evaluation of Training Devices," Human Factors, Vol. 21 (6) , pp. 711-720, 1979.
- Aikins, J.S., "Prototype and Production Rules: An Approach to Knowledge Representation for Hypothesis Formation," IJCAI, pp 1-3, 1979.
- Amarel, S., "Problems of Representation in Heuristic Problem Solving: Related Issues in the Development of Expert Systems," in "Methods of Heuristics" by Groner, R., Groner, M., and Bischof, W.F., (eds), Lea, 1983.
- Anderson, J.R., Cognitive Psychology and Its Implications, W.H. Freeman Co., 1980.
- Baddeley, A.D., The Psychology of Memory, Basic Books, Inc., New York, 1976.
- Bauer, M.A., "A Basis for the Acquisition of Procedures from Protocols", International Joint Conference on Artificial Intelligence, pp. 226-231, 1975.
- Baum, D. R., "A Framework and Topics for Empirical Research on Training Simulator Effectiveness," in Research Issues in the Determination of Simulator Fidelity : Proceedings of the ARI Sponsored Workshop, Hays (ed) US ARI for Behavioral and Social Sciences, Technical Report 547, Nov 1981.
- Baum, D. R., Riedel, S. L., and Hays, R. T., "Training Device Fidelity and Transfer of a Perceptual Motor Skill," Proceedings of 26th Human Factors Society Conference, pp. 746, 1982.
- Brown, J. S., "Applications of Artificial Intelligence in Maintenance Training," in New Concepts in Maintenance Trainers and Performance Aids, King, W. J. and Duva, J. S. (eds), NAVTRAEQUIPCEN IH-255, 1975.
- Caro, P. W., "Equipment-Device Task Commonality Analysis and Transfer of Training," HumRRO-TR-70-7, DAHC19-70-C-0012, June 1970.
- Chandrasekaran, B. and Mittal, "Deep versus Compiled Knowledge Approaches to Diagnostic Problem-solving," International Journal of Man-Machine Studies, Vol. 19, pp. 425-436, 1983.
- Clymer, A. B., "Simulation for Training and Decision-Making in Large-Scale Control Systems. Part I : Types of Training Simulators," Simulation, pp. 39, Aug 1980.
- Conover, W.J., and Iman, R.L., "On Some Alternative Procedures Using Ranks for the Analysis of Experimental Design," Communication in Statistics---Theory and Methods, Vol. A5, 1349-1368, 1976.

- Crawford, A. M., and Crawford, K. S., "Simulation of Operational Equipment with a Computer-Based Instructional System : A Low Cost Training Technology," Human Factors, Vol. 20, No. 2, April 1978.
- Curry, R. E., "A Model of Human Fault Detection for Complex Dynamic Process," in Human Detection and Diagnosis of System Failures, Rasmussen and Rouse, (eds) , Plenum, 1981.
- Davis, R., "Reasoning from First Principles in Electronic troubleshooting," International Journal of Man-Machine Studies, Vol. 19, pp. 403-423, 1983.
- de Kleer, J., "Causal and Teleological Reasoning in Circuit Recognition," PhD Thesis MIT, Sept 1979.
- de Kleer, J., and Brown, J.S., "Assumptions and Ambiguities in Mechanistic Mental Models," in Mental Models, edited by Dedre Gentner and Albert L. Stevens, Lawrence Erlbaum Associates Publishers, pp. 155-190, 1983.
- Duncan, K. D., and Shepherd, A., "A Simulator and Training Technique for Diagnosing Plant Failures from Control Panels," Ergonomics, Vol. 18, No. 6, pp. 627-641, 1975.
- Ericsson, K.A. and Simon H.A., Protocol Analysis: Verbal Reports as Data, The MIT Press, 1984.
- Fink, C. D., and Shriver, E. L., "Simulators for Maintenance Training : Some Issues, Problems and Areas for Future Research," Technical Report AFHRL-TR-78-27, Technical Training Division, Lowry AFB Colorado, July 1978.
- Finkelman, J. M., "Effect of Noise on Human Performance," Sound and Vibration, Vol. 36, pp. 26-28, 1975.
- Forbus, K., "A Study of Qualitative and Geometric Knowledge in Reasoning About Motion," Master 's Thesis, MIT, Feb 1980.
- Forbus, K., and Stevens, A., "Using Qualitative Simulation to Generate Explanations," Bolt, Beranek and Newman Inc. Report No. 4490, March 1981.
- Fowler, R. L., Williams, W. E., Fowler, M. G., and Young, D. D., "An Investigation of the Relationship Between Operator Performance and Operator Panel Layout for Continuous Tasks," Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Report No. AMRL-TR-68-170, 1968.
- Gagne, R. M., "Training Devices and Simulators : Some Research Issues," American Psychologist, Vol. 9, pp. 95-107, 1954.
- Genesereth, M.R., "Diagnosis Using Hierarchical Design Models,"

Proceedings of AAAI Conference, 1982.

Goodstein, L. P., "Summary of Workshop Discussions," in Human Detection and Diagnosis of System Failures, Rasmussen and Rouse, (eds) , Plenum, 1981.

Govindaraj, T., "Qualitative Simulation of Complex Dynamic Systems: An Application To A Marine Powerplant Under Supervisory Control," Proceedings of the 19th Annual Conference on Manual Control, Cambridge, MA, pp. 377-390, 23-25 May 1983.

Govindaraj, T., "Qualitative Modeling for Simulation of Large Dynamic Systems," Working Paper, Center for Man-Machine Systems Research, Georgia Institute of Technology, July 1984.

Hays, R. T., "Simulator Fidelity : A Concept Paper," US ARI for the Behavioral and Social Sciences, Technical Report No. 490, Nov 1980.

Hays, R. T., "Training Simulator Fidelity Guidance : The Iterative Data Base Approach," US ARI for the Behavioral and Social Sciences, Technical Report No. 545, Sept 1981.

Hays, R. T., "Research Issues in the Determination of Simulator Fidelity : Proceedings of the ARI Sponsored Workshop," US ARI for Behavioral and Social Sciences, Technical Report No. 547, Nov 1981.

Hunt, R. M. and Rouse, W. B., "Problem Solving Skills of Maintenance Trainees in Diagnosing Faults in Simulated Powerplants," Human Factors, 23(3), pp. 317-328 1981

Jaspers, B. W., and Hanley, G. S., "The Marine Engine Room Training Simulator at TNO Delft," in Ship Operation Automation, III, J. Vlietstra (ed) , North-Holland Co. IFIP, 1980.

Johnson, S. L., "Effect of Training Device on Retention and Transfer of a Procedural Task," Human Factors, Vol. 23 (3) , pp. 257-272, 1981.

Johnson, W.B., and Fath, J.L., "Implementation of a Mixed-Fidelity Approach to Maintenance Training," (TR-661), Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences, November 1984.

Johnson, W. B., and Rouse, W. B., "Training maintenance Technicians for Troubleshooting : Two Experiments with Computer Simulations," Human Factors, Vol. 24. No. 3, pp. 271-276, June 1982.

Kinkade, R. G., and Wheaton, G. R., "Training Device Design," in Human Engineering Guide to Equipment Design, Van Cott and Kinkade (eds) , ARI, 1972.

Kolodner, J.L., "Maintaining Organization in a Dynamic Long-Term Memory," Cognitive Science, Vol. 7, pp. 243-280, 1983.

- Kolodner, J.L., "Reconstructive Memory: A Computer Model," Cognitive Science, Vol 7(4), pp 281-328, 1983.
- Kolodner, J.L. and Kolodner, R.M., "An Algorithm for Diagnosis Based on Analysis of Previous Cases," in "Experience in Problem Solving: A Trilogy of Papers," GIT-ICS-84/16, School of Information and Computer Science, Georgia Institute of Technology, 1984.
- Kriessman, C., "A Spectrum of Simulators for Power Plant Training," Training and Development Journal, pp. 37-39, Sept 1981.
- Kuipers, B. and Kassirer, J.P., "Causal Reasoning in Medicine: Analysis of a Protocol," Cognitive Science, Vol 8(4), pp 363-386, 1984.
- Larkin, J., McDermott, J., Simon, D. P., and Simon, H. A., "Expert and Novice Performance in Solving Physics Problems," Science, Vol. 208, pp. 1335-1342, June 1980.
- Longman, M., Phelan, C. N., and Hansford, R. F., "Interactive Operation of Ship Simulators," MARSIM '81, 1981.
- Maddox, M.E., Johnson, W.B., and Frey, P.R., "A transfer of Training Study in the Nuclear Power Industry: From Simulation to Live Performance," The 29th Annual Conference of Human Factors Society, 1985.
- Martin, E. L., and Waag, W. L., "Contributions of Platform Motion to Simulator Effectiveness : Study 1-Basic Contact," Internal Report, AFHRL-TR-78-15, June 1978.
- McCallum, I. R., and Rawson, A. J., "The CARDIFF Ship Simulator : Design Features and Operational Philosophy," MARSIM '81, 1981.
- Mehle, T., "Hypothesis generation in an automobile malfunction inference task," Acta Psychologica, Vol. 52, pp. 87-116, 1982.
- Meishier, C. W., and Butler, G. J., "Air Combat Maneuvering Training in a Simulator," in Flight Simulation/Guidance System Simulation, Neuilly-Sur-Seine, France, NATO ARGARD Conference Proceedings No. 198, June 1976.
- Micheli, G. S., "Analysis of the Transfer of Training, Substitution, and Fidelity of Simulation of Transfer Equipment," TAEG Report No. 2, Training Analysis and Evaluation Group, NAVTRAEQUIPCEN Orlando FL, 1972.
- Miller, G.A., "The Magical Number Seven Plus or Minus Two: Some Limits on Our Capacity for Processing Information," Psychological Review, Vol. 63, pp. 81-9/, 1956.
- Miller, G. G., "Some Consideration in the Design and Utilization of Simulators for Technical Training," Technical Report : AFHRL-TR-

74-65, Aug 1974.

Miller, R. B., "Psychological Consideration in the Design of Training Equipment," WADC-TR-54-563, Wright Air Development Center, Wright-Patterson AFB, OH, Dec 1954.

Minsky, M., "A Framework for Representing Knowledge," in The Psychology of Computer Vision, edited by P. Winston, McGraw-Hill, New York, 1975.

Narva, M. A., "Formative Utilization of a Model for the Predictions of the Effectiveness of Training Devices," US ARI for Behavioral and Social Sciences, 16th Annual US Army Operations Research Symposium, Fort Lee, VA, Oct 1977.

Newell, A., and Simon, H.A., Human Problem Solving, Prentice-Hall, Inc., 1972.

Orlansky, J., "Interrelationships Between The Fidelity, Transfer and Cost Of Simulators: A Revised Look," Institute for Defense Analyses, Draft paper, Feb. 1984

Osgood, C. E., "The Similarity Paradox in Human Learning : A Resolution," Psychological Review, Vol. 56, pp. 132-143, 1947.

Pepler, R., "Thermal Comfort of Students in Climate Controlled and Non-climate Controlled Schools," ASHRAE Transaction, Vol. 78, pp. 97-109, 1972.

Polya, G., How to Solve It, Princeton University Press, 2nd printing, 1973.

Pople, H.E. Jr., "The Formation of Composite Hypotheses in Diagnostic Problem Solving An Exercise in Synthetic Reasoning," Proceeding of IJCAI, 1977.

Prophet, W. W., "The Importance of Training Requirements Information in the Design and Use of Aviation Training Devices," HumRRO, Professional Paper No. 8-66, Alexandria, VA, 1966.

Rasmussen, J., "On the Structure of Knowledge : A Morphology of Mental Models in a Man-Machine Systems Context," Riso National Laboratory, DK-4000, Roskilde, Denmark, Nov 1979.

Rasmussen, J., "The Human as a Systems Component," in Human Interaction with Computer, Smith, H.T. and Green, T.R.G. (eds), Academic Press, 1980.

Rasmussen, J., "The Role of Hierarchical Knowledge Representation in Decisionmaking and System Management," IEEE Transactions on SMC, Vol., CMS-15, No. 2, March/April 1985.

- Rouse, W. B., "Experimental Studies and Mathematical Models of Human Problem Solving Performance in Fault Diagnosis Tasks," in Human Detection and Diagnosis of System Failures, Rasmussen and Rouse (eds), Plenum, 1981.
- Rouse, W. B., "A Mixed-fidelity Approach to Technical Training," Journal of Educational Technology System, Vol. 11 (2), 1982-83.
- Rouse, W.B., "Models of Human Problem Solving: Detection, Diagnosis, and Compensation for System Failures," Automatica, Vol 19, no. 6, pp. 613-625, Nov 1983.
- Rouse, W.B., and Hunt, R.M., "Human Problem Solving in Fault Diagnosis Tasks," in Advances in Man-Machine Systems Research, edited by Rouse, W.B., JAI Press, Vol 1, pp. 195-222, 1984.
- Slovic, P., Fischhoff, B., and Lichtenstein, S., "Behavior Decision Theory", in Annual Review of Psychology, Vol 28, pp. 1-39, 1977.
- Schank, R.C., Dynamic Memory, Cambridge University Press, 1982.
- Semple, C. A., Hennesey, R. T., and Sanders, M. S., "Aircrew Training Devices : Fidelity Features," Canyon Research Group Inc. AFHRL-TR-80-36, Jan 1981.
- Shepherd, A., Marshall, E. C., Turner, A., and Duncan, K. D., "Diagnosis of Plant Failures from a Control Panel : A Comparison of Three Training Methods," Ergonomics, Vol. 20, No. 4, pp. 347-361, 1977.
- Sheridan, T., "Undersatnding human error and aiding human diagnostic behavior in nuclear power plants," in, Human Detection and Diagnosis of System Failures, edited by J. Rasmussen and W. Rouse, New York, Plenum Press, 1981.
- Shortliffe, E., Computer-based Medical Consultations: MYCIN, New York, American Elsevier, 1976.
- Spangenberg, R. W., "Selection of Simulation as an Instructional Medium," Proceedings Vol IV. Future of Simulators in Skills Training, First International Learning Technology Conference and Exposition on Applied Learning Technology. Society for Applied Learning Technology, pp.65, July 1976.
- Swezey, R.W., "Aspects of Criterion-Referenced Measurement in Performance Evaluation," Human Factors, 20(2), pp. 169-178, 1978.
- Szolovits, P. and Pauker, S.G., "Categorical and Probabilistic Reasoning in Medical Diagnosis," Artificial Intelligence, Vol 11, pp. 115-144, 1978.
- Tinker, M. A., "Illumination Intensity for Reading Newspaper Type," Journal of Education Psychology, Vol. 34, pp. 247-250, 1943.

- Trollip, S.R., "An Evaluation of a Complex Computer-Based Flight Procedures Trainer," Aviation Research Laboratory, Institute of Aviation, University of Illinois at Urbana-Champaign, Technical Report ARL-77-1/AFOSR-77-1, 1977.
- Tversky, A., and Kahneman, D., "Judgement under Uncertainty: Heuristics and Biases," Science, no. 185, pp. 1124-1131, 1974.
- Waterman, D.A., and Newell, A., "PAS-II: An Interactive Task-Free Version of an Automatic Protocol Analysis System," International Joint Conference on Artificial Intelligence, pp. 431-444, 1975.
- Wheaton, G. R., Rose, A. M., Fingerman, P. W., Korotkin, A. L., and Holding, D. H., "Evaluation of the Effectiveness of Training Device: Literature Review and Preliminary Model," Research Memorandum 76-6, US ARI for Behavioral and Social Sciences DAHC 19-73-C-0049, April 1976
- Williams, K. E., Goldberg, J., and D'Amico, A. D., "Transfer of training from Low to High Fidelity Simulators," CAORF 50-7919-02, Maritime Administration's Computer Aided Operations Research Facility, US Department of Commerce, Dec 1980.
- Williams, M.D., Hollan, J.D., and Stevens, A.L., "Human Reasoning About a Simple Physical System," in Mental Models, edited by Dedre Gentner and Albert L. Stevens, Lawrence Erlbaum Associates, New Jersey, 1983.

VITA

Yuan-Liang Su was born in Taiwan in 1951. He received the B.S. degree in Psychology from National Taiwan University, Taipei, Taiwan, in 1973, and the M.S. degree in Computer Science from Southern Illinois University, Edwardsville, Illinois, in 1979. He received a second M.S. degree in Operations Research in 1981, and the Ph.D. degree in Industrial and Systems Engineering in 1985, both from Georgia Institute of Technology, Atlanta, Georgia.

From 1982 to 1985, he worked as a research assistant at the Center for Man-Machine Systems Research, Georgia Institute of Technology, where he was engaged in research on modeling fault diagnosis performance of experts. He has recently joined the National Taiwan University, Taipei, Taiwan, as an assistant professor in the Department of Information Engineering. His research interests are in the areas of human problem solving, machine learning, artificial intelligence applications and expert systems.

Su, "Modeling Fault Diagnosis Performance"

Dr. Robert Ahlers
Code N711
Human Factors Laboratory
NAVTRAEQUIPCEN
Orlando, FL 32813

Dr. Ed Aiken
Navy Personnel R&D Center
San Diego, CA 92152

Dr. Robert Aiken
Temple University
School of Business Administration
Department of Computer and
Information Sciences
Philadelphia, PA 19122

Dr. James Algina
University of Florida
Gainesville, FL 32605

Dr. Earl A. Alluisi
HQ, AFHRL (AFSC)
Brooks AFB, TX 78235

Dr. Thomas H. Anderson
Center for the Study of Reading
174 Children's Research Center
51 Gerty Drive
Champaign, IL 61820

Technical Director, ARI
5001 Eisenhower Avenue
Alexandria, VA 22333

Special Assistant for Projects,
OASN(M&RA)
5D800, The Pentagon
Washington, DC 20350

Dr. Alan Baddeley
Medical Research Council
Applied Psychology Unit
15 Chaucer Road
Cambridge CB2 2EF
ENGLAND

Dr. Patricia Baggett
University of Colorado
Department of Psychology
Box 345
Boulder, CO 80309

Dr. Meryl S. Baker
Navy Personnel R&D Center
San Diego, CA 92152

Dr. John Black
Yale University
Box 11A, Yale Station
New Haven, CT 06520

Arthur S. Blaiwes
Code N711
Naval Training Equipment Center
Orlando, FL 32813

Dr. Robert Blanchard
Navy Personnel R&D Center
San Diego, CA 92152

Dr. Richard Braby
NTEC Code 10
Orlando, FL 32751

Dr. Robert Breaux
Code N-095R
NAVTRAEQUIPCEN
Orlando, FL 32813

Dr. Ann Brown
Center for the Study of Reading
University of Illinois
51 Gerty Drive
Champaign, IL 61820

Dr. Patricia A. Butler
NIE Mail Stop 1806
1200 19th St., NW
Washington, DC 20208

Dr. Robert Calfee
School of Education
Stanford University
Stanford, CA 94305

Dr. Joseph C. Campione
Center for the Study of Reading
University of Illinois
51 Gerty Drive
Champaign, IL 61820

Dr. Pat Carpenter
Carnegie-Mellon University
Department of Psychology
Pittsburgh, PA 15213

Su, "Modeling Fault Diagnosis Performance"

Dr. Robert Carroll
NAVOP 01B7
Washington, DC 20370

Dr. Fred Chang
Navy Personnel R&D Center
Code 51
San Diego, CA 92152

Dr. Davida Charney
Department of Psychology
Carnegie-Mellon University
Schenley Park
Pittsburgh, PA 15213

Dr. Stanley Collyer
Office of Naval Technology
Code 222
800 N. Quincy Street
Arlington, VA 22217-5000

Dr. Lee Cronbach
16 Laburnum Road
Atherton, CA 94205

Dr. Mary Cross
Department of Education
Adult Literacy Initiative
Room 4145
400 Maryland Avenue, SW
Washington, DC 20202

CTB/McGraw-Hill Library
2500 Garden Road
Monterey, CA 93940

CDR Mike Curran
Office of Naval Research
800 N. Quincy St.
Code 270
Arlington, VA 22217-5000

Bryan Dallman
AFHRL/LRT
Lowry AFB, CO 80230

Dr. Natalie John
Department of Computer and
Information Science
University of Oregon
Eugene, OR 97403

Defense Technical
Information Center
Cameron Station, Bldg 5
Alexandria, VA 22314
Attn: TC
(12 Copies)

Dr. Thomas M. Duffy
Communications Design Center
Carnegie-Mellon University
Schenley Park
Pittsburgh, PA 15213

Dr. Richard Duran
University of California
Santa Barbara, CA 93106

Edward E. Eddowes
CNATRA N301
Naval Air Station
Corpus Christi, TX 78419

Dr. John Ellis
Navy Personnel R&D Center
San Diego, CA 92252

Dr. Susan Embretson
University of Kansas
Psychology Department
Lawrence, KS 66045

Dr. Randy Engle
Department of Psychology
University of South Carolina
Columbia, SC 29208

Dr. William Epstein
University of Wisconsin
W. J. Brogden Psychology Bldg.
1202 W. Johnson Street
Madison, WI 53706

ERIC Facility-Acquisitions
4833 Rugby Avenue
Bethesda, MD 20014

Edward Esty
Department of Education, OERI
MS 40
1200 19th St., NW
Washington, DC 20208

Su, "Modeling Fault Diagnosis Performance"

Dr. Beatrice J. Farr
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Marshall J. Farr
2520 North Vernon Street
Arlington, VA 22207

Dr. Pat Federico
Code 511
NPRDC
San Diego, CA 92152

Mr. Wallace Feurzeig
Educational Technology
Bolt Beranek & Newman
10 Moulton St.
Cambridge, MA 02238

Dr. Linda Flower
Carnegie-Mellon University
Department of English
Pittsburgh, PA 15213

Dr. Carl H. Frederiksen
McGill University
3700 McTavish Street
Montreal, Quebec H3A 1Y2
CANADA

Dr. John R. Frederiksen
Bolt Beranek & Newman
50 Moulton Street
Cambridge, MA 02138

Dr. Arthur M. Glenberg
University of Wisconsin
W. J. Brogden Psychology Bldg.
1202 W. Johnson Street
Madison, WI 53706

Dr. Marvin D. Glock
13 Stone Hall
Cornell University
Ithaca, NY 14853

Dr. Sam Glucksberg
Princeton University
Department of Psychology
Green Hall
Princeton, NJ 08540

Dr. Joseph Goguen
Computer Science Laboratory
SRI International
333 Ravenswood Avenue
Menlo Park, CA 94025

Dr. Sherrie Gott
AFHRL/MODJ
Brooks AFB, TX 78235

Dr. Henry M. Halff
Halff Resources, Inc.
4918 33rd Road, North
Arlington, VA 22207

Mr. William Hartung
PEAM Product Manager
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Wayne Harvey
SRI International
333 Ravenswood Ave.
Room B-S324
Menlo Park, CA 94025

Prof. John R. Hayes
Carnegie-Mellon University
Department of Psychology
Schenley Park
Pittsburgh, PA 15213

Dr. Joan I. Heller
505 Haddon Road
Oakland, CA 94606

Dr. Melissa Holland
Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Dillon Inouye
WICAT Education Institute
Provo, UT 84057

Dr. Alice Isen
Department of Psychology
University of Maryland
Catonsville, MD 21228

Su, "Modeling Fault Diagnosis Performance"

Dr. Zachary Jacobson
Bureau of Management Consulting
365 Laurier Avenue West
Ottawa, Ontario K1A 0S5
CANADA

Dr. Claude Janvier
Directeur, CIRADE
Universite' du Quebec a Montreal
Montreal, Quebec H3C 3P8
CANADA

Margaret Jerome
c/o Dr. Peter Chandler
83, The Drive
Hove
Sussex
UNITED KINGDOM

Dr. Marcel Just
Carnegie-Mellon University
Department of Psychology
Schenley Park
Pittsburgh, PA 15213

Dr. Milton S. Katz
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. David Kieras
University of Michigan
Technical Communication
College of Engineering
1223 E. Engineering Building
Ann Arbor, MI 48109

Dr. Peter Kincaid
Training Analysis
& Evaluation Group
Department of the Navy
Orlando, FL 32813

Dr. Walter Kintsch
Department of Psychology
University of Colorado
Campus Box 345
Boulder, CO 80302

M. Diane Langston
Communications Design Center
Carnegie-Mellon University
Schenley Park
Pittsburgh, PA 15213

Dr. Charlotte Linde
SRI International
333 Ravenswood Avenue
Menlo Park, CA 94025

Dr. Arthur Melmed
U. S. Department of Education
724 Brown
Washington, DC 20208

Dr. George A. Miller
Department of Psychology
Green Hall
Princeton University
Princeton, NJ 08540

Dr. Lance A. Miller
IBM Thomas J. Watson
Research Center
P.O. Box 218
Yorktown Heights, NY 10598

Dr. Andrew R. Molnar
Scientific and Engineering
Personnel and Education
National Science Foundation
Washington, DC 20550

Dr. William Montague
NPRDC Code 13
San Diego, CA 92152

Assistant for Planning MANTRAPERS
NAVOP 01B6
Washington, DC 20370

Assistant for MPT Research,
Development and Studies
NAVOP 01B7
Washington, DC 20370

Director, Training Laboratory,
NPRDC (Code 05)
San Diego, CA 92152

Su, "Modeling Fault Diagnosis Performance"

Director, Manpower and Personnel
Laboratory,
NPRDC (Code 06)
San Diego, CA 92152

Director, Human Factors
& Organizational Systems Lab,
NPRDC (Code 07)
San Diego, CA 92152

Library, NPRDC
Code P201L
San Diego, CA 92152

Dr. Harry F. O'Neil, Jr.
Training Research Lab
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Director, Technology Programs,
Office of Naval Research
Code 200
800 North Quincy Street
Arlington, VA 22217-5000

Office of Naval Research,
Code 442PT
800 N. Quincy Street
Arlington, VA 22217-5000
(6 Copies)

Psychologist
Office of Naval Research
Branch Office, London
Box 39
FPO New York, NY 09510

Special Assistant for Marine
Corps Matters,
ONR Code 100M
800 N. Quincy St.
Arlington, VA 22217-5000

Psychologist
Office of Naval Research
Liaison Office, Far East
APO San Francisco, CA 96503

Dr. Judith Orasanu
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Douglas Pearse
DCIEM
Box 2000
Downsview, Ontario
CANADA

Dr. Nancy Pennington
University of Chicago
Graduate School of Business
1101 E. 58th St.
Chicago, IL 60637

Military Assistant for Training and
Personnel Technology,
OUSD (R & E)
Room 3D129, The Pentagon
Washington, DC 20301

Dr. Ray Perez
ARI (PERI-II)
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Steven Pinker
Department of Psychology
E10-018
M.I.T.
Cambridge, MA 02139

Dr. Tjeerd Plomp
Twente University of Technology
Department of Education
P.O. Box 217
7500 AE ENSCHEDE
THE NETHERLANDS

Dr. Martha Polson
Department of Psychology
Campus Box 346
University of Colorado
Boulder, CO 80309

Dr. Peter Polson
University of Colorado
Department of Psychology
Boulder, CO 80309

Dr. Steven E. Poltrock
MCC
9430 Research Blvd.
Echelon Bldg ■1
Austin, TX 78759-6509

Su, "Modeling Fault Diagnosis Performance"

Dr. Joseph Psotka
ATTN: PERI-1C
Army Research Institute
5001 Eisenhower Ave.
Alexandria, VA 22333

Dr. Lynne Reder
Department of Psychology
Carnegie-Mellon University
Schenley Park
Pittsburgh, PA 15213

Dr. Lauren Resnick
Learning R & D Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15213

Dr. Andrew M. Rose
American Institutes
for Research
1055 Thomas Jefferson St., NW
Washington, DC 20007

Dr. Robert Sasnor
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Ramsay W. Selden
NIE
Mail Stop 1241
1200 19th St., NW
Washington, DC 20208

Dr. Lee Shulman
Stanford University
1040 Cathcart Way
Stanford, CA 94305

Dr. Miriam Shustack
Code 51
Navy Personnel R & D Center
San Diego, CA 92152

Dr. Zita M Simutis
Instructional Technology
Systems Area
ARI
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. H. Wallace Sinaiko
Manpower Research
and Advisory Services
Smithsonian Institution
801 North Pitt Street
Alexandria, VA 22314

Dr. Edward E. Smith
Bolt Beranek & Newman, Inc.
50 Moulton Street
Cambridge, MA 02138

Dr. Robert Sternberg
Department of Psychology
Yale University
Box 11A, Yale Station
New Haven, CT 06520

Dr. Thomas Sticht
Navy Personnel R&D Center
San Diego, CA 92152

Dr. John Tangney
AFOSR/NL
Bolling AFB, DC 20332

Dr. Douglas Towne
Behavioral Technology Labs
1845 S. Elena Ave.
Redondo Beach, CA 90277

Dr. Paul Twohig
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Headquarters, U. S. Marine Corps
Code MPI-20
Washington, DC 20380

Dr. Beth Warren
Bolt Beranek & Newman, Inc.
50 Moulton Street
Cambridge, MA 02138

Dr. Barbara White
Bolt Beranek & Newman, Inc.
10 Moulton Street
Cambridge, MA 02238

Su, "Modeling Fault Diagnosis Performance"

Dr. Robert A. Wisher
U.S. Army Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Martin F. Wiskoff
Navy Personnel R & D Center
San Diego, CA 92152

Dr. Wallace Wulfeck, III
Navy Personnel R&D Center
San Diego, CA 92152

Dr. Joe Yasatuke
AFHRL/LRT
Lowry AFB, CO 80230

Dr. Joseph L. Young
Memory & Cognitive
Processes
National Science Foundation
Washington, DC 20550

END

FILMED

1-86

DTIC